


Math. Program., Ser. B (2018) 170:5–42  
<https://doi.org/10.1007/s10107-018-1287-z>



FULL LENGTH PAPER

# Bounds on entanglement dimensions and quantum graph parameters via noncommutative polynomial optimization

Sander Gribling<sup>1</sup> · David de Laat<sup>1</sup> ·  
Monique Laurent<sup>1,2</sup> 

Received: 8 January 2018 / Accepted: 2 May 2018 / Published online: 21 May 2018  
© The Author(s) 2018

**Abstract** In this paper we study optimization problems related to bipartite quantum correlations using techniques from tracial noncommutative polynomial optimization. First we consider the problem of finding the minimal entanglement dimension of such correlations. We construct a hierarchy of semidefinite programming lower bounds and show convergence to a new parameter: the minimal average entanglement dimension, which measures the amount of entanglement needed to reproduce a quantum correlation when access to shared randomness is free. Then we study optimization problems over synchronous quantum correlations arising from quantum graph parameters. We introduce semidefinite programming hierarchies and unify existing bounds on quantum chromatic and quantum stability numbers by placing them in the framework of tracial polynomial optimization.

**Keywords** Entanglement dimension · Polynomial optimization · Quantum graph parameters · Quantum correlation

**Mathematics Subject Classification** 05C15 · 81P40 · 81P45 · 90C22 · 90C26 · 90C30

---

The first and second authors are supported by the Netherlands Organization for Scientific Research, Grant Number 617.001.351, and the second author by the ERC Consolidator Grant QPROGRESS 615307.

---

✉ Monique Laurent  
[M.Laurent@cwi.nl](mailto:M.Laurent@cwi.nl)

<sup>1</sup> CWI, Amsterdam, Netherlands

<sup>2</sup> Tilburg University, Tilburg, Netherlands

# 1 Introduction

## 1.1 Bipartite quantum correlations

One of the distinguishing features of quantum mechanics is quantum entanglement, which allows for nonclassical correlations between spatially separated parties. In this paper we consider the problems of quantifying the advantage entanglement can bring (first investigated through Bell inequalities in the seminal work [3]) and quantifying the minimal amount of entanglement necessary for generating a given correlation (initiated in [5] and continued, e.g., in [43, 53, 60]).

Quantum entanglement has been widely studied in the bipartite correlation setting (for a survey, see, e.g., [44]). Here we have two parties, Alice and Bob, where Alice receives a question  $s$  taken from a finite set  $S$  and Bob receives a question  $t$  taken from a finite set  $T$ . The parties do not know each other's questions, and after receiving the questions they do not communicate. Then, according to some predetermined protocol, Alice returns an answer  $a$  from a finite set  $A$  and Bob returns an answer  $b$  from a finite set  $B$ . The probability that the parties answer  $(a, b)$  to questions  $(s, t)$  is given by a *bipartite correlation*  $P(a, b|s, t)$ , which satisfies  $P(a, b|s, t) \geq 0$  for all  $(a, b, s, t) \in \Gamma$  and  $\sum_{a,b} P(a, b|s, t) = 1$  for all  $(s, t) \in S \times T$ . We set  $\Gamma = A \times B \times S \times T$  throughout. Which bipartite correlations  $P = (P(a, b|s, t)) \in \mathbb{R}^\Gamma$  are possible depends on the additional resources available to the two parties Alice and Bob.

When the parties do not have access to additional resources the correlation  $P$  is *deterministic*, which means it is of the form  $P(a, b|s, t) = P_A(a|s) P_B(b|t)$  for all  $(a, b, s, t) \in \Gamma$ , where  $P_A = (P_A(a|s))$  and  $P_B = (P_B(b|t))$  take their values in  $\{0, 1\}$  and satisfy

$$\sum_a P_A(a|s) = \sum_b P_B(b|t) = 1 \quad \text{for all } (s, t) \in S \times T. \quad (1)$$

When the parties use local randomness the above functions  $P_A$  and  $P_B$  are convex combinations of 0/1-valued ones, that is,  $P_A$  and  $P_B$  take their values in  $[0, 1]$  and satisfy (1).

When the parties have access to shared randomness the resulting correlation  $P$  is a convex combination of deterministic correlations and  $P$  is said to be a *classical correlation*. The classical correlations form a polytope, denoted  $C_{loc}(\Gamma)$ , whose valid inequalities are known as *Bell inequalities* [3].

We are interested in the quantum setting, where the parties have access to a shared quantum state upon which they can perform measurements. The quantum setting can be modeled in different ways, leading to the so-called tensor model and commuting model; see the discussion, e.g., in [12, 36, 58].

In the *tensor model*, Alice and Bob each have access to “one half” of a finite dimensional *quantum state*, which is modeled by a unit vector  $\psi \in \mathbb{C}^d \otimes \mathbb{C}^d$  (for some  $d \in \mathbb{N}$ ). Alice and Bob determine their answers by performing a measurement on their part of the state. Such a measurement is modeled by a *positive operator valued measure* (POVM), which consists of a set of  $d \times d$  Hermitian positive semidefinite

matrices labeled by the possible answers and summing to the identity matrix. If Alice uses the POVM  $\{E_s^a\}_{a \in A}$  when she gets question  $s \in S$  and Bob uses the POVM  $\{F_t^b\}_{b \in B}$  when he gets question  $t \in T$ , then the probability of obtaining the answers  $(a, b)$  is given by

$$P(a, b|s, t) = \text{Tr}((E_s^a \otimes F_t^b)\psi\psi^*) = \psi^*(E_s^a \otimes F_t^b)\psi. \quad (2)$$

If the state  $\psi$  can be written as  $\psi = \psi_A \otimes \psi_B$ , then  $P(a, b|s, t) = (\psi_A^* E_s^a \psi_A)(\psi_B^* F_t^b \psi_B)$  for all  $(a, b, s, t)$ , and thus  $P$  is a classical correlation. Otherwise,  $\psi$  is said to be *entangled* and can be used to produce a nonclassical correlation  $P$ .

A correlation of the above form (2) is called a *quantum correlation*; it is said to be *realizable in the tensor model in local dimension  $d$*  (or in *dimension  $d^2$* ) when  $\psi \in \mathbb{C}^d \otimes \mathbb{C}^d$  and  $E_s^a, F_t^b \in \mathbb{C}^{d \times d}$ . Let  $C_q^d(\Gamma)$  be the set of such correlations and define

$$C_q(\Gamma) = \bigcup_{d \in \mathbb{N}} C_q^d(\Gamma).$$

Denote the smallest dimension needed to realize  $P \in C_q(\Gamma)$  in the tensor model by

$$D_q(P) = \min \{d^2 : d \in \mathbb{N}, P \in C_q^d(\Gamma)\}. \quad (3)$$

The set  $C_q^1(\Gamma)$  contains the deterministic correlations.<sup>1</sup> Hence, by Carathéodory's theorem,  $C_{loc}(\Gamma) \subseteq C_q^c(\Gamma)$  holds for  $c = |\Gamma| + 1 - |S||T|$ ; that is, quantum entanglement can be used as an alternative to shared randomness. If  $A, B, S$ , and  $T$  all contain at least two elements, then Bell [3] shows the inclusion  $C_{loc}(\Gamma) \subseteq C_q(\Gamma)$  is strict; that is, quantum entanglement can be used to obtain nonclassical correlations.

The second commonly used model to define quantum correlations is the *commuting model* (or *relativistic field theory model*). Here a correlation  $P \in \mathbb{R}^\Gamma$  is called a *commuting quantum correlation* if it is of the form

$$P(a, b|s, t) = \text{Tr}(X_s^a Y_t^b \psi\psi^*) = \psi^*(X_s^a Y_t^b)\psi, \quad (4)$$

where  $\{X_s^a\}_a$  and  $\{Y_t^b\}_b$  are POVMs consisting of bounded operators on a separable Hilbert space  $H$ , satisfying  $[X_s^a, Y_t^b] = X_s^a Y_t^b - Y_t^b X_s^a = 0$  for all  $(a, b, s, t) \in \Gamma$ , and where  $\psi$  is a unit vector in  $H$ . Such a correlation is said to be *realizable in dimension  $d = \dim(H)$  in the commuting model*. Denote the set of such correlations by  $C_{qc}^d(\Gamma)$  and set  $C_{qc}(\Gamma) = C_{qc}^\infty(\Gamma)$ . The smallest dimension needed to realize a quantum correlation  $P \in C_{qc}(\Gamma)$  is given by

$$D_{qc}(P) = \min \{d \in \mathbb{N} \cup \{\infty\} : P \in C_{qc}^d(\Gamma)\}. \quad (5)$$

If  $P \in C_q^d(\Gamma)$  has a decomposition (2) with  $d \times d$  matrices  $E_s^a, F_t^b$ , then  $P$  has a decomposition (4) with  $d^2 \times d^2$  matrices  $X_s^a = E_s^a \otimes I$  and  $Y_t^b = I \otimes F_t^b$ . This shows the inclusion  $C_q^d(\Gamma) \subseteq C_{qc}^{d^2}(\Gamma)$ , and thus

<sup>1</sup> In fact,  $C_q^1(\Gamma)$  consists of the correlations obtained using only local randomness.

$$D_{qc}(P) \leq D_q(P) \text{ for all } P \in C_q(\Gamma). \quad (6)$$

The minimum Hilbert space dimension in which a given quantum correlation  $P$  can be realized quantifies the minimal amount of entanglement needed to represent  $P$ . Computing  $D_q(P)$  is NP-hard [55], so a natural question is to find good lower bounds for the parameters  $D_q(P)$  and  $D_{qc}(P)$ . A main contribution of this paper is proposing a hierarchy of semidefinite programming lower bounds for these parameters.

As said above we have  $C_q^d(\Gamma) \subseteq C_{qc}^{d^2}(\Gamma)$ . Conversely, each finite dimensional commuting quantum correlation can be realized in the tensor model, although not necessarily in the same dimension [58] (see, e.g., [12] for a proof). This shows

$$C_q(\Gamma) = \bigcup_{d \in \mathbb{N}} C_q^d(\Gamma) \subseteq C_{qc}(\Gamma). \quad (7)$$

Using a direct sum construction one can show the sets  $C_q(\Gamma)$  and  $C_{qc}(\Gamma)$  are convex. Whether the two sets  $C_q(\Gamma)$  and  $C_{qc}(\Gamma)$  coincide is known as Tsirelson's problem.

In a recent breakthrough Slofstra [54] showed that the set  $C_q(\Gamma)$  is not closed for  $|A| \geq 8$ ,  $|B| \geq 2$ ,  $|S| \geq 184$ ,  $|T| \geq 235$ . More recently it was shown in [14] that the same holds for  $|A| \geq 2$ ,  $|B| \geq 2$ ,  $|S| \geq 5$ ,  $|T| \geq 5$ . Using a compactness argument one sees that the set  $C_q^d(\Gamma)$  is closed for all  $d$ . So, when  $C_q(\Gamma)$  is not closed, the inclusions  $C_q^d(\Gamma) \subset C_q(\Gamma)$  are all strict and there is a sequence  $\{P_i\} \subseteq C_q(\Gamma)$  with  $D_q(P_i) \rightarrow \infty$ . Moreover, since  $C_{qc}(\Gamma)$  is closed [15, Prop. 3.4], the inclusion  $C_q(\Gamma) \subseteq C_{qc}(\Gamma)$  is strict, thus settling Tsirelson's problem. Whether the closure of  $C_q(\Gamma)$  equals  $C_{qc}(\Gamma)$  for all  $\Gamma$  has been shown to be equivalent to having a positive answer to Connes' embedding conjecture in operator theory [21, 42]. This conjecture has been shown to have equivalent reformulations in many different fields; we refer to [23] for an algebraic reformulation in terms of trace positivity of noncommutative polynomials.

Further variations on the above definitions are possible. For instance, we can consider a mixed state  $\rho$  (a Hermitian positive semidefinite matrix  $\rho$  with  $\text{Tr}(\rho) = 1$ ) instead of a pure state  $\psi$ , where we replace the rank 1 matrix  $\psi\psi^*$  by  $\rho$  in the above definitions. By convexity this does not change the sets  $C_q(\Gamma)$  and  $C_{qc}(\Gamma)$ . It is shown in [53] that this also does not change the parameter  $D_q(P)$ , but it is unclear whether or not  $D_{qc}(P)$  might decrease. Another variation would be to use projection valued measures (PVMs) instead of POVMs, where the operators are projectors instead of positive semidefinite matrices. This again does not change the sets  $C_q(\Gamma)$  and  $C_{qc}(\Gamma)$  [40], but the dimension parameters can be larger when restricting to PVMs.

When the two parties have the same question sets ( $S = T$ ) and the same answer sets ( $A = B$ ), a bipartite correlation  $P \in \mathbb{R}^\Gamma$  is called *synchronous* if it satisfies

$$P(a, b|s, s) = 0 \quad \text{for all } s \in S \text{ and distinct } a, b \in A.$$

The sets of synchronous (commuting) quantum correlations are denoted  $C_{q,s}(\Gamma)$  and  $C_{qc,s}(\Gamma)$ , respectively. We have  $C_{q,s}(\Gamma) \subseteq C_{qc,s}(\Gamma)$  and the set  $C_{qc,s}(\Gamma)$  is closed. The synchronous correlation sets are already rich enough in the sense that it is still the case that Connes' embedding conjecture holds if and only if  $\text{cl}(C_{q,s}(\Gamma)) = C_{qc,s}(\Gamma)$

for all  $\Gamma$  [13, Thm. 3.7]. The quantum graph parameters discussed in Sect. 1.3 will be defined through optimization problems over synchronous quantum correlations.

For a synchronous quantum correlation  $P$  it turns out that its local dimension  $D_q(P)$  is given by the factorization rank of an associated completely positive semidefinite matrix  $M_P$ . Recall that a matrix  $M \in \mathbb{R}^{n \times n}$  is called *completely positive semidefinite* if there exist  $d \in \mathbb{N}$  and  $d \times d$  Hermitian positive semidefinite matrices  $X_1, \dots, X_n$  with  $M = (\text{Tr}(X_i X_j))$ . The minimal such  $d$  is its *completely positive semidefinite rank*, denoted  $\text{cpsd-rank}(M)$ . Completely positive semidefinite matrices are used in [28] to model quantum graph parameters and the  $\text{cpsd-rank}$  is investigated in [16, 17, 49, 50]. In Sect. 2 we show the following link between synchronous quantum correlations and  $\text{cpsd-rank}$ .

**Proposition 1** *The smallest local dimension in which a synchronous quantum correlation  $P$  can be realized is given by the completely positive semidefinite rank of the matrix  $M_P$  indexed by  $S \times A$  with entries  $(M_P)_{(s,a),(t,b)} = P(a, b|s, t)$  for  $(a, b, s, t) \in \Gamma$ . That is,  $D_q(P) = \text{cpsd-rank}(M_P)$ .*

In [16] we used techniques from tracial polynomial optimization to define a semidefinite programming hierarchy  $\{\xi_r^{\text{cpsd}}(M)\}_{r \in \mathbb{N}}$  of lower bounds on  $\text{cpsd-rank}(M)$ . By the above result this hierarchy gives lower bounds on the smallest local dimension in which a synchronous correlation can be realized in the tensor model. However, as shown in [16], this hierarchy typically does not converge to  $\text{cpsd-rank}(M)$  but instead (under a certain flatness condition) to a parameter  $\xi_*^{\text{cpsd}}(M)$ , which can be seen as a block-diagonal version of the completely positive semidefinite rank. This *flatness condition* is a rank stabilization condition on the optimal solution of the semidefinite program defining  $\xi_r^{\text{cpsd}}(M)$ ; for a formal definition see (21) in Sect. 3.3.

Here we use similar techniques, now exploiting the special structure of quantum correlations, to construct a hierarchy  $\{\xi_r^q(P)\}$  of lower bounds on the minimal dimension  $D_q(P)$  of any—not necessarily synchronous—quantum correlation  $P$ . The hierarchy converges (under flatness) to a parameter  $\xi_*^q(P)$ , and using the additional structure we can show that  $\xi_*^q(P)$  is equal to an interesting parameter  $A_q(P) \leq D_q(P)$ . This parameter describes the minimal average entanglement dimension of a correlation when the parties have free access to shared randomness; see Sect. 1.2.

In the rest of the introduction we give a road map through the contents of the paper and state the main results. We will introduce the necessary background along the way.

## 1.2 A hierarchy for the average entanglement dimension

We give here an overview of the results in Sect. 3 about bounding the entanglement dimension of general (non synchronous) correlations. We are interested in the minimal entanglement dimension needed to realize a given correlation  $P \in C_q(\Gamma)$ . If  $P$  is deterministic or only uses local randomness, then  $D_q(P) = D_{qc}(P) = 1$ . But other classical correlations (which use shared randomness) have  $D_q(P) \geq D_{qc}(P) > 1$ , which means the shared quantum state is used as a shared randomness resource. In [5] the concept of dimension witness is introduced, where a  $d$ -dimensional witness is defined as a halfspace containing  $\text{conv}(C_q^d(\Gamma))$ , but not the full set  $C_q(\Gamma)$ . As a measure of entanglement this suggests the parameter

$$\inf \left\{ \max_{i \in [I]} D_q(P_i) : I \in \mathbb{N}, \lambda \in \mathbb{R}_+^I, \sum_{i=1}^I \lambda_i = 1, P = \sum_{i=1}^I \lambda_i P_i, P_i \in C_q(\Gamma) \right\}. \quad (8)$$

Observe that, for a bipartite correlation  $P$ , this parameter is equal to 1 if and only if  $P$  is classical. Hence, it more closely measures the minimal entanglement dimension when the parties have free access to shared randomness. From an operational point of view, (8) can be interpreted as follows. Before the game starts the parties select a finite number of pure states  $\psi_i$  ( $i \in I$ ) (instead of a single one), in possibly different dimensions  $d_i$ , and POVMs  $\{E_s^a(i)\}_a, \{F_t^b(i)\}_b$  for each  $i \in I$  and  $(s, t) \in S \times T$ . As before, we assume that the parties cannot communicate after receiving their questions  $(s, t)$ , but now they do have access to shared randomness, which they use to decide on which state  $\psi_i$  to use. The parties proceed to measure state  $\psi_i$  using POVMs  $\{E_s^a(i)\}_a, \{F_t^b(i)\}_b$ , so that the probability of answers  $(a, b)$  is given by the quantum correlation  $P_i$ . Equation (8) then asks for the largest dimension needed in order to generate  $P$  when access to shared randomness is free.

It is not clear how to compute (8). Here we propose a variation of (8), and we provide a hierarchy of semidefinite programs that converges to it under flatness. Instead of considering the largest dimension needed to generate  $P$ , we consider the *average* dimension. That is, we minimize  $\sum_{i \in I} \lambda_i D_q(P_i)$  over all convex combinations  $P = \sum_{i \in I} \lambda_i P_i$ . Hence, the *minimal average entanglement dimension* is defined by

$$A_q(P) = \inf \left\{ \sum_{i=1}^I \lambda_i D_q(P_i) : I \in \mathbb{N}, \lambda \in \mathbb{R}_+^I, \sum_{i=1}^I \lambda_i = 1, P = \sum_{i=1}^I \lambda_i P_i, P_i \in C_q(\Gamma) \right\} \quad (9)$$

in the tensor model. In the commuting model, the parameter  $A_{qc}(P)$  is defined by the same expression with  $D_q(P_i)$  being replaced by  $D_{qc}(P_i)$ . Observe that we need not replace  $C_q(\Gamma)$  by  $C_{qc}(\Gamma)$  since  $D_{qc}(P) = \infty$  for any  $P \in C_{qc}(\Gamma) \setminus C_q(\Gamma)$ . Moreover, in view of (6), we have the inequality

$$A_{qc}(P) \leq A_q(P) \quad \text{for all } P \in C_q(\Gamma). \quad (10)$$

It follows by convexity that for the above definitions it does not matter whether we use pure or mixed states. We show that for the average minimal entanglement dimension it also does not matter whether we use the tensor or commuting model.

**Proposition 2** *For any  $P \in C_q(\Gamma)$  we have  $A_q(P) = A_{qc}(P)$ .*

We have  $A_q(P) \leq D_q(P)$  and  $A_{qc}(P) \leq D_{qc}(P)$  for  $P \in C_q(\Gamma)$ , with equality if  $P$  is an extreme point of  $C_q(\Gamma)$ . Hence, we have  $D_q(P) = D_{qc}(P)$  if  $P$  is an extreme point of  $C_q(\Gamma)$ . We show that the parameter  $A_q(P)$  can be used to distinguish between classical and nonclassical correlations.

**Proposition 3** *For  $P \in C_q(\Gamma)$  we have  $A_q(P) = 1$  if and only if  $P \in C_{loc}(\Gamma)$ .*

As mentioned before, there exist sets  $\Gamma$  for which  $C_q(\Gamma)$  is not closed [14, 54], which implies the existence of a sequence  $\{P_i\} \subseteq C_q(\Gamma)$  such that  $D_q(P) \rightarrow \infty$ . We show this also implies the existence of such a sequence with  $A_q(P_i) \rightarrow \infty$ .

**Proposition 4** *If  $C_q(\Gamma)$  is not closed, then there exists a sequence  $\{P_i\} \subseteq C_q(\Gamma)$  with  $A_q(P_i) \rightarrow \infty$ .*

Using tracial polynomial optimization we construct a hierarchy  $\{\xi_r^q(P)\}$  of lower bounds on  $A_{qc}(P)$ . For each  $r \in \mathbb{N}$  this is a semidefinite program, and for  $r = \infty$  it is an infinite dimensional semidefinite program. We further define a (hyperfinite) variation  $\xi_*^q(P)$  of  $\xi_\infty^q(P)$  by adding a finite rank constraint on the matrix variable, so that

$$\xi_1^q(P) \leq \xi_2^q(P) \leq \dots \leq \xi_\infty^q(P) \leq \xi_*^q(P) \leq A_{qc}(P).$$

We do not know whether  $\xi_\infty^q(P) = \xi_*^q(P)$  always holds. First we show that we imposed enough constraints in the bounds  $\xi_r^q(P)$  so that  $\xi_*^q(P) = A_{qc}(P)$ .

**Proposition 5** *For any  $P \in C_q(\Gamma)$  we have  $\xi_*^q(P) = A_{qc}(P)$ .*

Then we show that the infinite dimensional semidefinite program  $\xi_\infty^q(P)$  is the limit of the finite dimensional semidefinite programs.

**Proposition 6** *For any  $P \in C_q(\Gamma)$  we have  $\xi_r^q(P) \rightarrow \xi_\infty^q(P)$  as  $r \rightarrow \infty$ .*

Finally we give a flatness criterion under which finite convergence  $\xi_r^q(P) = \xi_*^q(P)$  holds. The definition of flatness follows later in the paper [see (21)]; here we only note that it is a rank stabilization property which is easy to check given a solution to  $\xi_r^q(P)$ .

**Proposition 7** *If  $\xi_r^q(P)$  admits a  $(\lceil r/3 \rceil + 1)$ -flat optimal solution, then we have  $\xi_r^q(P) = \xi_*^q(P)$ .*

### 1.3 Quantum graph parameters

Nonlocal games have been introduced in quantum information theory as abstract models to quantify the power of entanglement, in particular, in how much the sets  $C_q(\Gamma)$  and  $C_{qc}(\Gamma)$  differ from  $C_{loc}(\Gamma)$ . A *nonlocal game* is defined by a probability distribution  $\pi: S \times T \rightarrow [0, 1]$  and a predicate  $f: A \times B \times S \times T \rightarrow \{0, 1\}$ . Alice and Bob receive a question pair  $(s, t) \in S \times T$  with probability  $\pi(s, t)$ . They know the game parameters  $\pi$  and  $f$ , but they do not know each other's questions, and they cannot communicate after they receive their questions. Their answers  $(a, b)$  are determined according to some correlation  $P \in \mathbb{R}^\Gamma$ , called their *strategy*, on which they may agree before the start of the game, and which can be classical or quantum depending on whether  $P$  belongs to  $C_{loc}(\Gamma)$ ,  $C_q(\Gamma)$ , or  $C_{qc}(\Gamma)$ . Then their corresponding winning probability is given by

$$\sum_{(s,t) \in S \times T} \pi(s, t) \sum_{(a,b) \in A \times B} P(a, b|s, t) f(a, b, s, t). \quad (11)$$

A strategy  $P$  is called *perfect* if the above winning probability is equal to one, that is, if for all  $(a, b, s, t) \in \Gamma$  we have



$$(\pi(s, t) > 0 \quad \text{and} \quad f(a, b, s, t) = 0) \implies P(a, b|s, t) = 0. \quad (12)$$

Computing the maximum winning probability of a nonlocal game is an instance of linear optimization [of the function (11)] over  $C_{loc}(\Gamma)$  in the classical setting, and over  $C_q(\Gamma)$  or  $C_{qc}(\Gamma)$  in the quantum setting. Since the inclusion  $C_{loc}(\Gamma) \subseteq C_q(\Gamma)$  can be strict, the maximum winning probability can be higher when the parties have access to entanglement; see the CHSH game [10]. In fact there are nonlocal games that can be won with probability 1 by using entanglement, but with probability strictly less than 1 in the classical setting; see the Mermin-Peres magic square game [34, 47].

The quantum graph parameters are analogues of the classical parameters defined through the coloring and stability number games as described below. These nonlocal games use the set  $[k]$  (whose elements are denoted as  $a, b$ ) and the set  $V$  of vertices of a graph  $G$  (whose elements are denoted as  $i, j$ ) as question and answer sets.

In the *quantum coloring game*, introduced in [1, 9], we have a graph  $G = (V, E)$  and an integer  $k$ . Here we have question sets  $S = T = V$  and answer sets  $A = B = [k]$ , and the distribution  $\pi$  is strictly positive on  $V \times V$ . The predicate  $f$  is such that the players' answers have to be consistent with having a  $k$ -coloring of  $G$ ; that is,  $f(a, b, i, j) = 0$  precisely when  $(i = j \text{ and } a \neq b)$  or  $(\{i, j\} \in E \text{ and } a = b)$ . This expresses the fact that if Alice and Bob receive the same vertex, they should return the same color and if they receive adjacent vertices, they should return distinct colors. A perfect classical strategy exists if and only if a perfect deterministic strategy exists, and a perfect deterministic strategy corresponds to a  $k$ -coloring of  $G$ . Hence the smallest number  $k$  of colors for which there exists a perfect classical strategy is equal to the classical chromatic number  $\chi(G)$ . It is therefore natural to define the quantum chromatic number as the smallest  $k$  for which there exists a perfect quantum strategy. Observe that such a strategy is necessarily synchronous since, in view of (12),  $f(a, b, i, i) = 0$  when  $a \neq b$  implies  $P(a, b|i, i) = 0$  when  $a \neq b$ .

**Definition 1** The (commuting) quantum chromatic number  $\chi_q(G)$  (resp.,  $\chi_{qc}(G)$ ) is the smallest  $k \in \mathbb{N}$  for which there exists a synchronous correlation  $P = (P(a, b|i, j))$  in  $C_{q,s}([k]^2 \times V^2)$  (resp.,  $C_{qc,s}([k]^2 \times V^2)$ ) such that

$$P(a, a|i, j) = 0 \quad \text{for all} \quad a \in [k], \{i, j\} \in E.$$

In the *quantum stability number game*, introduced in [32, 51], we again have a graph  $G = (V, E)$  and  $k \in \mathbb{N}$ , but now we use the question set  $[k] \times [k]$  and the answer set  $V \times V$ . The distribution  $\pi$  is again strictly positive on the question set and now the predicate  $f$  of the game is such that the players' answers have to be consistent with having a stable set of size  $k$ , that is,  $f(i, j, a, b) = 0$  precisely when  $(a = b \text{ and } i \neq j)$  or  $[a \neq b \text{ and } (i = j \text{ or } \{i, j\} \in E)]$ . This expresses the fact that when Alice and Bob receive the same index  $a = b \in [k]$ , they should answer with the same vertex  $i = j$  of  $G$ , and if they receive distinct indices  $a \neq b$  from  $[k]$ , they should answer with distinct nonadjacent vertices  $i$  and  $j$  of  $G$ . There is a perfect classical strategy precisely when there exists a stable set of size  $k$ , so that the largest integer  $k$  for which there exists a perfect classical strategy is equal to the stability number  $\alpha(G)$ . Again, such a strategy is necessarily synchronous, so we get the following definition.



**Definition 2** The (commuting) stability number  $\alpha_q(G)$  (resp.,  $\alpha_{qc}(G)$ ) is the largest integer  $k \in \mathbb{N}$  for which there exists a synchronous correlation  $P = (P(i, j|a, b))$  in  $C_{q,s}(V^2 \times [k]^2)$  (resp.,  $C_{qc,s}(V^2 \times [k]^2)$ ) such that

$$P(i, j|a, b) = 0 \quad \text{whenever} \quad (i = j \text{ or } \{i, j\} \in E) \text{ and } a \neq b \in [k].$$

The classical parameters  $\chi(G)$  and  $\alpha(G)$  are NP-hard. The same holds for the quantum coloring number  $\chi_q(G)$  [20], and also for the quantum stability number  $\alpha_q(G)$  in view of the following reduction to coloring shown in [32]:

$$\chi_q(G) = \min \{k \in \mathbb{N} : \alpha_q(G \square K_k) = |V|\}. \quad (13)$$

Here  $G \square K_k$  is the Cartesian product of the graph  $G = (V, E)$  and the complete graph  $K_k$ . By construction we have

$$\chi_{qc}(G) \leq \chi_q(G) \leq \chi(G) \quad \text{and} \quad \alpha(G) \leq \alpha_q(G) \leq \alpha_{qc}(G).$$

The separations between  $\chi_q(G)$  and  $\chi(G)$ , and between  $\alpha_q(G)$  and  $\alpha(G)$ , can be exponentially large in the number of vertices. This is the case for the graphs with vertex set  $\{\pm 1\}^n$  for  $n$  a multiple of 4, where two vertices are adjacent if they are orthogonal [1, 32, 33]. It is well known that the chromatic number of a graph increases by 1 if we add a new vertex that is adjacent to all other vertices. Surprisingly, this is not true in general for the quantum chromatic number [31]. While it was recently shown that the sets  $C_{q,s}(\Gamma)$  and  $C_{qc,s}(\Gamma)$  can be different [14], it is not known whether there is a separation between the parameters  $\chi_q(G)$  and  $\chi_{qc}(G)$ , and between  $\alpha_q(G)$  and  $\alpha_{qc}(G)$ .

We now give an overview of the results of Sect. 4 and refer to that section for formal definitions. In Sect. 4.1 we first reformulate the quantum graph parameters in terms of  $C^*$ -algebras, which allows us to use techniques from tracial polynomial optimization to formulate bounds on the quantum graph parameters. We define a hierarchy  $\{\gamma_r^{\text{col}}(G)\}$  of lower bounds on the commuting quantum chromatic number and a hierarchy  $\{\gamma_r^{\text{stab}}(G)\}$  of upper bounds on the commuting quantum stability number. We show the following convergence results for these hierarchies.

**Proposition 8** *There is an  $r_0 \in \mathbb{N}$  such that  $\gamma_r^{\text{col}}(G) = \chi_{qc}(G)$  and  $\gamma_r^{\text{stab}}(G) = \alpha_{qc}(G)$  for all  $r \geq r_0$ . Moreover, if  $\gamma_r^{\text{col}}(G)$  admits a flat optimal solution, then  $\gamma_r^{\text{col}}(G) = \chi_q(G)$ , and if  $\gamma_r^{\text{stab}}(G)$  admits a flat optimal solution, then  $\gamma_r^{\text{stab}}(G) = \alpha_q(G)$ .*

Then in Sect. 4.2 we define tracial analogues  $\{\xi_r^{\text{stab}}(G)\}$  and  $\{\xi_r^{\text{col}}(G)\}$  of Lasserre type bounds on  $\alpha(G)$  and  $\chi(G)$  that provide hierarchies of bounds for their quantum analogues. These bounds are more economical than the bounds  $\gamma_r^{\text{col}}(G)$  and  $\gamma_r^{\text{stab}}(G)$  (since they use less variables) and they also permit to recover some known bounds for the quantum parameters. We show that  $\xi_*^{\text{stab}}(G)$ , which is the parameter  $\xi_\infty^{\text{stab}}(G)$  with an additional rank constraint on the matrix variable, coincides with the projective packing number  $\alpha_p(G)$  from [51] and that  $\xi_\infty^{\text{stab}}(G)$  upper bounds  $\alpha_{qc}(G)$ .

**Proposition 9** *We have  $\xi_*^{\text{stab}}(G) = \alpha_p(G) \geq \alpha_q(G)$  and  $\xi_\infty^{\text{stab}}(G) \geq \alpha_{qc}(G)$ .*

Next, we consider the chromatic number. The tracial hierarchy  $\{\xi_r^{\text{col}}(G)\}$  unifies two known bounds: the projective rank  $\xi_f(G)$ , a lower bound on the quantum chromatic number from [32], and the tracial rank  $\xi_{tr}(G)$ , a lower bound on the commuting quantum chromatic number from [46]. In [13, Cor. 3.10] it is shown that the projective rank and the tracial rank coincide if Connes' embedding conjecture is true.

**Proposition 11** *We have  $\xi_*^{\text{col}}(G) = \xi_f(G) \leq \chi_q(G)$  and  $\xi_\infty^{\text{col}}(G) = \xi_{tr}(G) \leq \chi_{qc}(G)$ .*

We compare the hierarchies  $\xi_r^{\text{col}}(G)$  and  $\gamma_r^{\text{col}}(G)$ , and the hierarchies  $\xi_r^{\text{stab}}(G)$  and  $\gamma_r^{\text{stab}}(G)$ . For the coloring parameters, we show the analogue of reduction (13).

**Proposition 12** *For  $r \in \mathbb{N} \cup \{\infty\}$  we have  $\gamma_r^{\text{col}}(G) = \min\{k : \xi_r^{\text{stab}}(G \square K_k) = |V|\}$ .*

We show an analogous statement for the stability parameters, when using the homomorphic graph product of  $K_k$  with the complement of  $G$ , denoted here as  $K_k \star G$ , and the following reduction shown in [32]:

$$\alpha_q(G) = \max\{k \in \mathbb{N} : \alpha_q(K_k \star G) = k\}.$$

**Proposition 13** *For  $r \in \mathbb{N} \cup \{\infty\}$  we have  $\gamma_r^{\text{stab}}(G) = \max\{k : \xi_r^{\text{stab}}(K_k \star G) = k\}$ .*

Finally, we show that the hierarchies  $\{\gamma_r^{\text{col}}(G)\}$  and  $\{\gamma_r^{\text{stab}}(G)\}$  refine the hierarchies  $\{\xi_r^{\text{col}}(G)\}$  and  $\{\xi_r^{\text{stab}}(G)\}$ .

**Proposition 14** *For  $r \in \mathbb{N} \cup \{\infty, *\}$ ,  $\xi_r^{\text{col}}(G) \leq \gamma_r^{\text{col}}(G)$  and  $\xi_r^{\text{stab}}(G) \geq \gamma_r^{\text{stab}}(G)$ .*

## 1.4 Techniques from noncommutative polynomial optimization

In a (commutative) polynomial optimization problem we minimize a multivariate polynomial  $f(x_1, \dots, x_n)$  over a feasible region defined by polynomial inequalities. Such a problem has the form

$$\inf\{f(x_1, \dots, x_n) : x \in \mathbb{R}^n, g(x_1, \dots, x_n) \geq 0 \text{ for } g \in \mathcal{G}\}$$

for some finite set  $\mathcal{G}$  of multivariate polynomials. Lasserre [24] and Parrilo [45] introduced the moment/sum-of-squares method to solve such problems (see, e.g., [25, 27] for details). The moment method is based on the observation that the above polynomial optimization problem is equivalent to minimizing  $\int f(x) d\mu(x)$  over all probability measures  $\mu$  supported on the set  $D(\mathcal{G}) = \{x \in \mathbb{R}^n : g(x) \geq 0 \text{ for } g \in \mathcal{G}\}$ . In turn, this is equivalent to minimizing  $L(f)$  over all linear functionals  $L$  on the space of polynomials satisfying  $L(p) \geq 0$  for all polynomials  $p$  that are nonnegative on  $D(\mathcal{G})$ . To get a tractable relaxation we then consider the linear functionals  $L$  on the space of polynomials up to degree  $2r$  and require that  $L$  is nonnegative on all squares  $s^2$  and weighted squares  $s^2g$  (for  $g \in \mathcal{G}$ ) of degree at most  $2r$ . This condition can be expressed with a polynomially sized semidefinite program for any fixed  $r$ . These

relaxations are good in the sense that, under a mild assumption,<sup>2</sup> they converge to the optimal value of the polynomial optimization problem as  $r$  goes to infinity.

In [37, 48] this approach has been extended to the general eigenvalue optimization problem, which is a problem of the form

$$\inf \left\{ \psi^* f(X_1, \dots, X_n) \psi : d \in \mathbb{N}, \psi \in \mathbb{C}^d \text{ unit vector}, X_1, \dots, X_n \in \mathbb{C}^{d \times d}, \right. \\ \left. g(X_1, \dots, X_n) \geq 0 \text{ for } g \in \mathcal{G} \right\}.$$

Here, the matrix variables  $X_i$  are allowed to have any dimension  $d \in \mathbb{N}$  and  $\{f\} \cup \mathcal{G}$  is a set of symmetric polynomials in noncommutative variables. In a tracial optimization problem, instead of minimizing the smallest eigenvalue of  $f(X_1, \dots, X_n)$ , we minimize its normalized trace  $\text{Tr}(f(X_1, \dots, X_n))/d$  (so that the identity matrix has trace one) [6–8, 22]. Such a problem has the form

$$\inf \left\{ \text{Tr}(f(X_1, \dots, X_n))/d : d \in \mathbb{N}, X_1, \dots, X_n \in \mathbb{C}^{d \times d}, \right. \\ \left. g(X_1, \dots, X_n) \geq 0 \text{ for } g \in \mathcal{G} \right\},$$

where the matrix variables  $X_i$  may again have any dimension  $d$  and  $\{f\} \cup \mathcal{G}$  is a set of symmetric polynomials in noncommutative variables. The moment approach for these two problems again relies on minimizing  $L(f)$ , where  $L$  is a linear functional on the space of *noncommutative* polynomials that either models  $\psi^* f(X_1, \dots, X_n) \psi$  or models the *normalized trace evaluation*  $\text{Tr}(f(X_1, \dots, X_n))/d$ .

Let us focus on the tracial setting which is the setting used in this paper. As in the commutative case, one obtains tractable (semidefinite programming) relaxations by requiring  $L$  to “behave like a trace evaluation on noncommutative polynomials of degree at most  $2r$ ”. Specifically, we ask  $L$  to be nonnegative on all Hermitian squares  $s^*s$  and weighted Hermitian squares  $s^*gs$  (for  $g \in \mathcal{G}$ ) of degree at most  $2r$ , and we require the new *tracial* condition  $L(pq) = L(qp)$ , which indeed holds for trace evaluations; see Sect. 3.3 for details. Under an analogous mild assumption, the asymptotic limit of these relaxations is well understood: we obtain a solution  $(X_1, \dots, X_n)$  living in a  $C^*$ -algebra  $\mathcal{A}$  equipped with a tracial state  $\tau$ . The question thus becomes: when can such a solution be converted into a solution to the original tracial optimization problem, i.e., to a solution living in a usual matrix algebra?

For our purposes, a  $C^*$ -algebra  $\mathcal{A}$  can be defined as a norm closed  $*$ -subalgebra of the space  $\mathcal{B}(H)$  of bounded operators on a complex Hilbert space  $H$ . Here, the involution  $*$  on  $\mathcal{B}(H)$  is the usual adjoint operation, and a  $*$ -subalgebra is an algebra that is closed under taking adjoints. When  $H$  has finite dimension  $d$  this means  $\mathcal{A}$  is a *matrix  $*$ -algebra*, i.e.,  $\mathcal{A}$  is a subalgebra of  $\mathbb{C}^{d \times d}$  that is closed under taking complex conjugates. Examples of matrix  $*$ -algebras include the full matrix algebra  $\mathbb{C}^{d \times d}$  or the  $*$ -algebra generated by given matrices  $X_1, \dots, X_n \in \mathbb{C}^{d \times d}$ , denoted  $\mathbb{C}\langle X_1, \dots, X_n \rangle$ . An algebra is called *finite dimensional* if it is finite dimensional as a vector space.

<sup>2</sup> To be precise, it suffices that there exists a constant  $R > 0$  such that  $R - \sum_i x_i^2$  can be written as a sum of weighted squares  $s^2g$  with  $g \in \mathcal{G} \cup \{1\}$ . This is called the Archimedean condition of the quadratic module associated to  $\mathcal{G}$ , see Sect. 3.3.

Essential for understanding the asymptotic limit of the above relaxations for tracial polynomial optimization are the following results due to Artin and Wedderburn (see [2, 59]): Any finite dimensional  $C^*$ -algebra is  $(*)$ -isomorphic to a matrix  $*$ -algebra containing the identity, and in turn any such matrix  $*$ -algebra is isomorphic to a direct sum of full matrix algebras. We record the latter result for future reference:

**Theorem 1** ([2, 59]) *Let  $\mathcal{A}$  be a complex matrix  $*$ -subalgebra of  $\mathbb{C}^{d \times d}$  containing the identity. Then there exists a unitary matrix  $U$  and integers  $K, m_k, n_k$  for  $k \in [K]$  such that*

$$U \mathcal{A} U^* = \bigoplus_{k=1}^K (\mathbb{C}^{n_k \times n_k} \otimes I_{m_k}) \quad \text{and} \quad d = \sum_{k=1}^K m_k n_k.$$

Going back to the question above about the asymptotic limit of the relaxations to the tracial optimization problem: when the obtained solution  $(X_1, \dots, X_n)$  lives in a finite dimensional  $C^*$ -algebra it can be converted into an optimal matrix solution to the original tracial optimization problem. As we will later see (Theorem 3) this happens when the limit linear functional  $L$  satisfies some finite rank condition since then  $L$  is a convex combination of trace evaluations at matrix tuples  $(X_1, \dots, X_n)$  satisfying  $g(X_1, \dots, X_n) \geq 0$  for all  $g \in \mathcal{G}$ . In addition note that this may happen at a relaxation of finite order  $r$  when the optimal solution  $L$  satisfies the so-called flatness condition (see Theorems 3 and 4).

An important feature in noncommutative polynomial optimization is the dimension independence: the optimization is over all possible matrix sizes  $d \in \mathbb{N}$ . In fact, this was the original motivation in the works [36] and [12], where noncommutative polynomial optimization was first used for approximating the set  $C_{qc}(\Gamma)$  of commuting quantum correlations and the maximum winning probability of nonlocal games over  $C_{qc}(\Gamma)$  (and, more generally, for computing Bell inequality violations). In some applications one may want to restrict to optimizing over matrices with restricted size  $d$ . In [35, 38] techniques are developed that allow to incorporate this dimension restriction by suitably selecting the linear functionals  $L$  in a specified space; this is used to give bounds on the maximum violation of a Bell inequality in a fixed dimension. A related natural problem is to decide what is the minimum dimension  $d$  needed to realize a given algebraically defined object, such as a (commuting) quantum correlation  $P$ . Here we propose an approach based on tracial polynomial optimization: starting from the observation that the trace of the  $d \times d$  identity matrix gives its size  $d$ , we consider the problem of minimizing  $L(1)$  where  $L$  is a linear functional now modeling the *non-normalized* matrix trace. This approach has been used in several recent works [16, 39, 57] for lower bounding factorization ranks of matrices and tensors.

## 2 Entanglement dimension of synchronous quantum correlations

By combining the proofs from [52] (see also [32]) and [46] one can show the following link between the minimum local dimension of a synchronous correlation and the

completely positive semidefinite rank of an associated completely positive semidefinite matrix.

**Proposition 1** *The smallest local dimension in which a synchronous quantum correlation  $P$  can be realized is given by the completely positive semidefinite rank of the matrix  $M_P$  indexed by  $S \times A$  with entries  $(M_P)_{(s,a),(t,b)} = P(a, b|s, t)$  for  $(a, b, s, t) \in \Gamma$ . That is,  $D_q(P) = \text{cpsd-rank}(M_P)$ .*

*Proof* Suppose first that  $(\psi, E_s^a, F_t^b)$  is a realization of  $P$  in local dimension  $d$  as in (2). We will show  $M_P$  is completely positive semidefinite with  $\text{cpsd-rank}_{\mathbb{C}}(M_P) \leq d$ .

Taking the Schmidt decomposition of  $\psi$ , there exist nonnegative scalars  $\{\lambda_i\}$  and orthonormal bases  $\{u_i\}$  and  $\{v_i\}$  of  $\mathbb{C}^d$  such that  $\psi = \sum_{i=1}^d \sqrt{\lambda_i} u_i \otimes v_i$ .<sup>3</sup> If we replace  $\psi$  by  $\sum_{i=1}^d \sqrt{\lambda_i} v_i \otimes v_i$  and  $E_s^a$  by  $U E_s^a U^*$ , where  $U$  is the unitary matrix for which  $u_i = U v_i$  for all  $i$ , then  $(\sum_{i=1}^d \sqrt{\lambda_i} u_i \otimes v_i, E_s^a, F_t^b)$  still realizes  $P$  and is of the same dimension  $d$ .

Given such a realization  $(\sum_{i=1}^d \sqrt{\lambda_i} v_i \otimes v_i, E_s^a, F_t^b)$  of  $P$ , we define the matrices

$$K = \sum_{i=1}^d \sqrt{\lambda_i} v_i v_i^*, \quad X_s^a = K^{1/2} E_s^a K^{1/2}, \quad Y_t^b = K^{1/2} F_t^b K^{1/2}.$$

By using the identities  $\text{vec}(K) = \psi$  and

$$\text{vec}(K)^*(E_s^a \otimes F_t^b)\text{vec}(K) = \text{Tr}(K E_s^a K F_t^b) = \text{Tr}(K^{1/2} E_s^a K^{1/2} K^{1/2} F_t^b K^{1/2}), \quad (14)$$

and substituting  $X_s^a = K^{1/2} E_s^a K^{1/2}$  and  $Y_t^b = K^{1/2} F_t^b K^{1/2}$ , we see that

$$P(a, b|s, t) = \langle X_s^a, Y_t^b \rangle \quad \text{for all } a, b, s, t, \quad (15)$$

and

$$\langle K, K \rangle = 1, \quad \sum_a X_s^a = \sum_b Y_t^b = K \quad \text{for all } s, t. \quad (16)$$

For any  $s \in S$ , as  $P$  is synchronous we have  $1 = \sum_{a,b} P(a, b|s, s) = \sum_a P(a, a|s, s)$ . Then the Cauchy–Schwarz inequality gives

$$\begin{aligned} 1 &= \sum_a P(a, a|s, s) = \sum_a \langle X_s^a, Y_s^a \rangle \leq \sum_a \langle X_s^a, X_s^a \rangle^{1/2} \langle Y_s^a, Y_s^a \rangle^{1/2} \\ &\leq \left( \sum_a \langle X_s^a, X_s^a \rangle \right)^{1/2} \left( \sum_a \langle Y_s^a, Y_s^a \rangle \right)^{1/2} \\ &\leq \left\langle \sum_a X_s^a, \sum_a X_s^a \right\rangle^{1/2} \left\langle \sum_a Y_s^a, \sum_a Y_s^a \right\rangle^{1/2} = \langle K, K \rangle = 1. \end{aligned}$$

<sup>3</sup> The Schmidt decomposition  $\psi = \sum_{i=1}^d \sqrt{\lambda_i} u_i \otimes v_i$  of  $\psi \in \mathbb{C}^d \otimes \mathbb{C}^d$  can be viewed as the singular value decomposition  $\sum_{i=1}^d \sqrt{\lambda_i} u_i v_i^*$  of the matrix  $\text{vec}^{-1}(\psi)$ , where  $\text{vec} : \mathbb{C}^{d \times d} \rightarrow \mathbb{C}^d \otimes \mathbb{C}^d$  is the operation that sends  $uv^*$  to  $u \otimes v$ .

Thus all inequalities above are equalities. The first inequality being an equality shows that there exist  $\alpha_{s,a} \geq 0$  such that  $X_s^a = \alpha_{s,a} Y_s^a$  for all  $a, s$ . The second inequality being an equality shows that there exist  $\beta_s$  such that  $\|X_s^a\| = \beta_s \|Y_s^a\|$  for all  $a, s$ . Hence,

$$\beta_s \|Y_s^a\| = \|X_s^a\| = \|\alpha_{s,a} Y_s^a\| = \alpha_{s,a} \|Y_s^a\| = \alpha_{s,a} \|Y_s^a\| \quad \text{for all } a, s,$$

which shows  $X_s^a = \beta_s Y_s^a$  for all  $a, s$ . Since  $\sum_a X_s^a = K = \sum_a Y_s^a$ , we have  $\beta_s = 1$  for all  $s$ . Thus  $X_s^a = Y_s^a$  for all  $a, s$ . Therefore,

$$(M_P)_{(s,a),(t,b)} = \langle X_s^a, X_t^b \rangle \quad \text{for all } a, b, s, t,$$

which shows  $M_P$  is completely positive semidefinite with  $\text{cpsd-rank}_{\mathbb{C}}(M_P) \leq d$ .

For the other direction we suppose  $\{X_s^a\}$  are smallest possible Hermitian positive semidefinite matrices such that  $(M_P)_{(s,a),(t,b)} = \langle X_s^a, X_t^b \rangle$  for all  $a, s, t, b$ . Then,

$$1 = \sum_{a,b} P(a, b|s, t) = \sum_{a,b} \langle X_s^a, X_t^b \rangle = \left\langle \sum_a X_s^a, \sum_b X_t^b \right\rangle \quad \text{for all } s, t,$$

which shows the existence of a matrix  $K$  such that  $K = \sum_a X_s^a$  for all  $s$ . We have  $\langle K, K \rangle = 1$  and thus  $\text{vec}(K)$  is a unit vector. Moreover, since the factorization of  $M_P$  is chosen of smallest possible size, the matrix  $K$  is invertible. Set  $E_s^a = K^{-1/2} X_s^a K^{-1/2}$  for all  $s, a$ , so that  $\sum_a E_s^a = I$  for all  $s$ . Then, using again (14) we obtain

$$P(a, b|s, t) = (M_P)_{(s,a),(t,b)} = \langle X_s^a, X_t^b \rangle = \text{vec}(K)^* (E_s^a \otimes E_t^b) \text{vec}(K),$$

which shows  $P$  has a realization of local dimension  $\text{cpsd-rank}_{\mathbb{C}}(M_P)$ .  $\square$

### 3 A hierarchy for the minimal entanglement dimension

#### 3.1 The minimal average entanglement dimension

Here we investigate some properties of the average entanglement dimension  $A_q(\cdot)$ , which was introduced in Sect. 1.2 in (9). We start by showing that it does not matter whether we use the tensor model or the commuting model.

**Proposition 2** *For any  $P \in C_q(\Gamma)$  we have  $A_q(P) = A_{qc}(P)$ .*

*Proof* The inequality  $A_{qc}(P) \leq A_q(P)$  was observed in (10). For the reverse inequality assume we have a decomposition  $P = \sum_{i=1}^I \lambda_i P_i$ , which is feasible for  $A_{qc}(P)$ . This means we have POVMs  $\{X_s^a(i)\}_a$  and  $\{Y_t^b(i)\}_b$  in  $\mathbb{C}^{d_i \times d_i}$  with  $[X_s^a(i), Y_t^b(i)] = 0$  and unit vectors  $\psi_i \in \mathbb{C}^{d_i}$  such that  $P_i(a, b|s, t) = \psi_i^* X_s^a(i) Y_t^b(i) \psi_i$  for all  $(a, b, s, t) \in \Gamma$  and  $i \in [I]$ . We will construct another decomposition of  $P$  which will provide a feasible solution to  $A_q(P)$  with value at most  $\sum_i \lambda_i d_i$ .

Fix some index  $i \in [I]$ . Applying Theorem 1 to the matrix  $\ast$ -algebra  $\mathbb{C}\langle\{X_s^a(i)\}_{a,s}\rangle$  generated by the matrices  $X_s^a(i)$  for  $(a, s) \in A \times S$  shows that there exist a unitary matrix  $U_i$  and integers<sup>4</sup>  $K_i, m_k, n_k$  such that

$$U_i \mathbb{C}\langle\{X_s^a(i)\}_{a,s}\rangle U_i^* = \bigoplus_{k=1}^{K_i} (\mathbb{C}^{n_k \times n_k} \otimes I_{m_k}) \quad \text{and} \quad d_i = \sum_{k=1}^{K_i} m_k n_k.$$

By assumption each matrix  $Y_t^b(i)$  commutes with all the matrices in  $\mathbb{C}\langle\{X_s^a(i)\}_{a,s}\rangle$ , and thus  $U_i Y_t^b(i) U_i^*$  lies in the algebra  $\bigoplus_k (I_{n_k} \otimes \mathbb{C}^{m_k \times m_k})$ . Hence, we may assume

$$X_s^a(i) = \bigoplus_{k=1}^{K_i} E_s^a(i, k) \otimes I_{m_k}, \quad Y_t^b(i) = \bigoplus_{k=1}^{K_i} I_{n_k} \otimes F_t^b(i, k), \quad \psi_i = \bigoplus_{k=1}^{K_i} \psi_{i,k},$$

with  $E_s^a(i, k) \in \mathbb{C}^{n_k \times n_k}$ ,  $F_t^b(i, k) \in \mathbb{C}^{m_k \times m_k}$ , and  $\psi_{i,k} \in \mathbb{C}^{n_k} \otimes \mathbb{C}^{m_k}$ . Then we have

$$\begin{aligned} P_i(a, b|s, t) &= \text{Tr}(X_s^a(i) Y_t^b(i) \psi_i \psi_i^*) \\ &= \sum_{k=1}^{K_i} \|\psi_{i,k}\|^2 \underbrace{\text{Tr} \left( E_s^a(i, k) \otimes F_t^b(i, k) \frac{\psi_{i,k} \psi_{i,k}^*}{\|\psi_{i,k}\|^2} \right)}_{Q_{i,k}(a, b|s, t)}, \end{aligned}$$

where  $Q_{i,k} \in C_q(\Gamma)$ . As  $\sum_k \|\psi_{i,k}\|^2 = \|\psi_i\|^2 = 1$ , we have that  $P_i = \sum_k \|\psi_{i,k}\|^2 Q_{i,k}$  is a convex combination of the  $Q_{i,k}$ 's.

We now show that  $Q_{i,k} \in C_q^{\min\{m_k, n_k\}}(\Gamma)$ . Consider the Schmidt decomposition  $\psi_{i,k} / \|\psi_{i,k}\| = \sum_{l=1}^{\min\{m_k, n_k\}} \lambda_{i,k,l} v_{i,k,l} \otimes w_{i,k,l}$ , where  $\lambda_{i,k,l} \geq 0$  and  $\{v_{i,k,l}\}_{l=1}^{n_k} \subseteq \mathbb{C}^{n_k}$  and  $\{w_{i,k,l}\}_{l=1}^{m_k} \subseteq \mathbb{C}^{m_k}$  are orthonormal bases. Define unitary matrices  $V_k \in \mathbb{C}^{n_k \times n_k}$  and  $W_k \in \mathbb{C}^{m_k \times m_k}$  such that  $V_k v_{i,k,l}$  is the  $l$ th unit vector in  $\mathbb{R}^{n_k}$  for  $1 \leq l \leq n_k$  and  $W_k w_{i,k,l}$  is the  $l$ th unit vector in  $\mathbb{R}^{m_k}$  for  $1 \leq l \leq m_k$ . Let  $E_s^a(i, k)'$  (resp.,  $F_t^b(i, k)'$ ) be the leading principal submatrices of  $V_k E_s^a(i, k) V_k^*$  (resp.,  $W_k F_t^b(i, k) W_k^*$ ) of size  $\min\{m_k, n_k\}$ . Moreover, set  $\phi_{i,k} = \sum_{l=1}^{\min\{m_k, n_k\}} \lambda_{i,k,l} e_l \otimes e_l$ , where  $e_l$  is the  $l$ th unit vector in  $\mathbb{R}^{\min\{m_k, n_k\}}$ . Then we have

$$\begin{aligned} Q_{i,k}(a, b|s, t) &= \text{Tr} \left( E_s^a(i, k) \otimes F_t^b(i, k) \frac{\psi_{i,k} \psi_{i,k}^*}{\|\psi_{i,k}\|^2} \right) \\ &= \sum_{l, l'=1}^{\min\{m_k, n_k\}} \lambda_{i,k,l} \lambda_{i,k,l'} (v_{i,k,l}^* E_s^a(i, k) v_{i,k,l'}) (w_{i,k,l}^* F_t^b(i, k) w_{i,k,l'}) \end{aligned}$$

<sup>4</sup> We omit the explicit dependence on  $i$  in the integers  $m_k, n_k$  to simplify the notation.



$$\begin{aligned}
&= \sum_{l,l'=1}^{\min\{m_k, n_k\}} \lambda_{i,k,l} \lambda_{i,k,l'} (e_l^* E_s^a(i, k)' e_{l'}) (e_l^* F_t^b(i, k)' e_{l'}) \\
&= \text{Tr}((E_s^a(i, k)' \otimes F_t^b(i, k)') \phi_{i,k} \phi_{i,k}^*),
\end{aligned}$$

which shows  $Q_{i,k} \in C_q^{\min\{m_k, n_k\}}(\Gamma)$ .

Combining the convex decompositions  $P = \sum_i \lambda_i P_i$  and  $P_i = \sum_k \|\psi_{i,k}\|^2 Q_{i,k}$ , we get the following convex decomposition  $P = \sum_{i,k} \lambda_i \|\psi_{i,k}\|^2 Q_{i,k}$ , from which we obtain

$$\begin{aligned}
A_q(P) &\leq \sum_{i,k} \lambda_i \|\psi_{i,k}\|^2 \min\{m_k, n_k\}^2 \leq \sum_{i,k} \lambda_i \min\{m_k, n_k\}^2 \\
&\leq \sum_{i,k} \lambda_i m_k n_k = \sum_i \lambda_i d_i. \quad \square
\end{aligned}$$

We now show that the parameter  $A_q(\cdot)$  permits to characterize classical correlations.

**Proposition 3** *For  $P \in C_q(\Gamma)$  we have  $A_q(P) = 1$  if and only if  $P \in C_{loc}(\Gamma)$ .*

*Proof* If  $P \in C_{loc}(\Gamma)$ , then  $P$  can be written as a convex combination of deterministic correlations (which belong to  $C_q^1(\Gamma)$ ), and thus  $A_q(P) = 1$ .

For the reverse implication, assume  $A_q(P) = 1$ . Then there exist a sequence of convex decompositions  $P = \sum_{i \in I^l} \lambda_i^l P_i^l$  indexed by  $l \in \mathbb{N}$ , with  $\{P_i^l\} \subseteq C_q(\Gamma)$  and  $\lim_{l \rightarrow \infty} \sum_{i \in I^l} \lambda_i^l D_q(P_i^l) = 1$ . Decompose the set  $I^l$  as the disjoint union  $I_-^l \cup I_+^l$ , where  $D_q(P_i^l) = 1$  for  $i \in I_-^l$  and  $D_q(P_i^l) > 1$  for  $i \in I_+^l$ . Let  $\varepsilon > 0$ . Then, for all  $l$  sufficiently large we have

$$\begin{aligned}
1 + \sum_{i \in I_+^l} \lambda_i &= \left(1 - \sum_{i \in I_+^l} \lambda_i^l\right) + 2 \sum_{i \in I_+^l} \lambda_i^l \leq \sum_{i \in I_-^l} \lambda_i^l + \sum_{i \in I_+^l} \lambda_i^l D_q(P_i^l) \\
&= \sum_{i \in I^l} \lambda_i^l D_q(P_i^l) \leq 1 + \varepsilon,
\end{aligned}$$

implying  $\sum_{i \in I_+^l} \lambda_i^l \leq \varepsilon$ . This shows that the sequence  $\mu^l := \sum_{i \in I_-^l} \lambda_i^l$  tends to 1 as  $l \rightarrow \infty$ . The correlation  $P^l := \sum_{i \in I_-^l} \lambda_i^l P_i^l / \mu^l$  is a convex combination of deterministic correlations and thus it belongs to  $C_{loc}(\Gamma)$ . Moreover,  $P^l \rightarrow P$  as  $l \rightarrow \infty$ , which implies  $P \in C_{loc}(\Gamma)$ .  $\square$

As we already observed earlier, when the set  $C_q(\Gamma)$  is not closed, the inclusion  $C_q^d(\Gamma) \subseteq C_q(\Gamma)$  is strict for all  $d$  (because with a compactness argument one can show that  $C_q^d(\Gamma)$  is closed), and thus there exists a sequence  $\{P_i\} \subseteq C_q(\Gamma)$  with  $D_q(P_i) \rightarrow \infty$  as  $i \rightarrow \infty$ . We show the analogous unboundedness property for the average entanglement dimension  $A_q(\cdot)$ . For the proof we will use the fact that also the sets  $C_{qc}^d(\Gamma)$  are closed for all  $d \in \mathbb{N}$ .

**Proposition 4** *If  $C_q(\Gamma)$  is not closed, then there exists a sequence  $\{P_i\} \subseteq C_q(\Gamma)$  with  $A_q(P_i) \rightarrow \infty$ .*

*Proof* Assume for contradiction there exists an integer  $K$  such that  $A_q(P) \leq K$  for all  $P \in C_q(\Gamma)$ . We will show this results in a uniform upper bound  $K'$  on  $D_{qc}(P)$ , which, in view of (7), implies that  $C_q(\Gamma)$  is equal to the closed set  $C_{qc}^{K'}(\Gamma)$ , contradicting the assumption that  $C_q(\Gamma)$  is not closed. For this, we will first show that any  $P \in C_q(\Gamma)$  belongs to  $\text{conv}(C_{qc}^K(\Gamma))$ .

In a first step observe that any  $P \in C_q(\Gamma) \setminus \text{conv}(C_{qc}^K(\Gamma))$  can be decomposed as

$$P = \mu_1 R_1 + (1 - \mu_1) Q_1, \quad (17)$$

where  $R_1 \in C_q(\Gamma)$ ,  $Q_1 \in \text{conv}(C_{qc}^K(\Gamma))$ , and  $0 < \mu_1 \leq K/(K+1)$ . Indeed, by assumption and using Proposition 2,  $A_{qc}(P) = A_q(P) \leq K$ , so  $P$  can be written as a convex combination  $P = \sum_{i \in I} \lambda_i P_i$  with  $\{P_i\} \subseteq C_q(\Gamma)$  and  $\sum_{i \in I} \lambda_i D_{qc}(P_i) \leq K$ . As  $P \notin \text{conv}(C_{qc}^K(\Gamma))$ , the set  $J$  of indices  $i \in I$  with  $D_{qc}(P_i) \geq K+1$  is non empty. Then  $(K+1) \sum_{i \in J} \lambda_i \leq \sum_{i \in J} \lambda_i D_{qc}(P_i) \leq K$ , and thus  $0 < \mu_1 := \sum_{i \in J} \lambda_i \leq K/(K+1)$ . Hence (17) holds after setting  $R_1 = (\sum_{i \in J} \lambda_i P_i)/\mu_1$  and  $Q_1 = (\sum_{i \in I \setminus J} \lambda_i P_i)/(1 - \mu_1)$ .

As  $R_1 \in C_q(\Gamma) \setminus \text{conv}(C_{qc}^K(\Gamma))$ , we may repeat the same argument for  $R_1$ . By iterating we obtain for each integer  $k \in \mathbb{N}$  a decomposition

$$\begin{aligned} P &= \mu_1 \mu_2 \dots \mu_k R_k \\ &\quad + \underbrace{(1 - \mu_1) Q_1 + \mu_1 (1 - \mu_2) Q_2 + \dots + \mu_1 \mu_2 \dots \mu_{k-1} (1 - \mu_k) Q_k}_{=(1 - \mu_1 \mu_2 \dots \mu_k) \hat{Q}_k}, \end{aligned}$$

where  $R_k \in C_q(\Gamma)$ ,  $\hat{Q}_k \in \text{conv}(C_{qc}^K(\Gamma))$  and  $\mu_1 \mu_2 \dots \mu_k \leq (K/(K+1))^k$ . Then the sequence  $\mu_1 \mu_2 \dots \mu_k$  tends to 0 as  $k \rightarrow \infty$ . As the entries of  $R_k$  lie in  $[0, 1]$  we can conclude that  $\mu_1 \mu_2 \dots \mu_k R_k$  tends to 0 as  $k \rightarrow \infty$ . Hence the sequence  $(\hat{Q}_k)_k$  has a limit  $\hat{Q}$  and  $P = \hat{Q}$  holds. As all  $\hat{Q}_k$  lie in the compact set  $\text{conv}(C_{qc}^K(\Gamma))$ , we also have  $P \in \text{conv}(C_{qc}^K(\Gamma))$ . So we reach a contradiction, which shows  $C_q(\Gamma) \subseteq \text{conv}(C_{qc}^K(\Gamma))$ .

The extreme points of the compact convex set  $\text{conv}(C_{qc}^K(\Gamma))$  lie in  $C_{qc}^K(\Gamma)$ , so, by the Carathéodory theorem, any  $P \in \text{conv}(C_{qc}^K(\Gamma))$  is a convex combination of  $c$  elements from  $C_{qc}^K(\Gamma)$ , where  $c = |\Gamma| + 1 - |S||T|$ . By using a direct sum construction one can obtain  $D_{qc}(P) \leq cK$ , which shows  $K' := cK$  is a uniform upper bound on  $D_{qc}(P)$  for all  $P \in C_q(\Gamma)$ .  $\square$

### 3.2 Setup of the hierarchy

We will now construct a hierarchy of lower bounds on the minimal entanglement dimension, using its formulation via  $A_{qc}(\cdot)$ . Our approach is based on noncommutative polynomial optimization, thus similar to the approach we used in [16] for bounding matrix factorization ranks.

We first need some notation. Set  $\mathbf{x} = \{x_s^a : (a, s) \in A \times S\}$  and  $\mathbf{y} = \{y_t^b : (b, t) \in B \times T\}$ , and let  $\langle \mathbf{x}, \mathbf{y}, z \rangle_r$  be the set of all words in the  $n = |S||A| + |T||B| + 1$  symbols  $x_s^a$ ,  $y_t^b$ , and  $z$ , having length at most  $r$ . Moreover, set  $\langle \mathbf{x}, \mathbf{y}, z \rangle = \langle \mathbf{x}, \mathbf{y}, z \rangle_\infty$ . We equip  $\langle \mathbf{x}, \mathbf{y}, z \rangle_r$  with an involution  $w \mapsto w^*$  that reverses the order of the symbols in the words and leaves the symbols  $x_s^a$ ,  $y_t^b$ ,  $z$  invariant; e.g.,  $(x_s^a z)^* = z x_s^a$ . Let  $\mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_r$  be the vector space of all real linear combinations of the words of length (aka degree) at most  $r$ . The space  $\mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle = \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_\infty$  is the  $*$ -algebra with Hermitian generators  $\{x_s^a\}$ ,  $\{y_t^b\}$ , and  $z$ , and the elements in this algebra are called *noncommutative polynomials* in the variables  $\{x_s^a\}$ ,  $\{y_t^b\}$ ,  $z$ .

The hierarchy of bounds on  $A_{qc}(P)$  is based on the following idea: For any feasible solution to  $A_{qc}(P)$ , its objective value can be modeled as  $L(1)$  for a certain tracial linear form  $L$  on the space of noncommutative polynomials (truncated to degree  $2r$ ).

Indeed, assume  $\{(P_i, \lambda_i)_i\}$  is a feasible solution to the program defining  $A_{qc}(P)$  (introduced in Sect. 1.2). That is,  $P = \sum_i \lambda_i P_i$  with  $\lambda_i \geq 0$ ,  $\sum_i \lambda_i = 1$  and  $P_i \in C_q(\Gamma)$ . Assume  $P_i(a, b|s, t) = \text{Tr}(X_s^a(i)Y_t^b(i)\psi_i\psi_i^*)$  where  $\psi \in \mathbb{C}^{d_i}$  and the POVM's  $\{X_s^a(i)\}$ ,  $\{Y_t^b(i)\} \subset \mathbb{C}^{d_i \times d_i}$  are as in (4), that is, for all  $(s, t, a, b) \in \Gamma$  the matrices  $X_s^a(i)$  and  $Y_t^b(i)$  commute:  $[X_s^a(i), Y_t^b(i)] = X_s^a(i)Y_t^b(i) - Y_t^b(i)X_s^a(i) = 0$ . For  $r \in \mathbb{N} \cup \{\infty\}$ , consider the linear functional  $L \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_{2r}^*$  defined by

$$L(p) = \sum_i \lambda_i \text{Re}(\text{Tr}(p(\mathbf{X}(i), \mathbf{Y}(i), \psi_i\psi_i^*))) \quad \text{for } p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_{2r}.$$

Here, for each index  $i$ , we set

$$\mathbf{X}(i) = (X_s^a(i) : (a, s) \in A \times S), \quad \mathbf{Y}(i) = (Y_t^b(i) : (b, t) \in B \times T),$$

and we replace the variables  $x_s^a$ ,  $y_t^b$ ,  $z$  by  $X_s^a(i)$ ,  $Y_t^b(i)$ , and  $\psi_i\psi_i^*$ , respectively. First note that we have  $L(1) = \sum_i \lambda_i d_i$ . That is,  $L(1)$  is equal to the objective value of the feasible solution  $\{(P_i, \lambda_i)_i\}$  to  $A_{qc}(P)$ . Secondly, for all  $(s, t, a, b) \in \Gamma$  we have  $L(x_s^a y_t^b z) = P(a, b|s, t)$ .

We will now identify several computationally tractable properties that this linear functional  $L$  satisfies. The hierarchy of lower bounds on  $A_{qc}(P)$  then consists of optimization problems where we minimize  $L(1)$  over the set of linear functionals that satisfy these properties.

First note that  $L$  is *symmetric*, that is,  $L(w) = L(w^*)$  for all  $w \in \langle \mathbf{x}, \mathbf{y}, z \rangle_{2r}$ , and *tracial*, that is,  $L(w w') = L(w' w)$  for all  $w, w' \in \langle \mathbf{x}, \mathbf{y}, z \rangle$  with  $\deg(w w') \leq 2r$ .

Next, for all  $p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_{r-1}$  we have

$$L(p^* x_s^a p) = \sum_i \lambda_i \text{Re}(\text{Tr}(C(i)^* X_s^a(i) C(i))) \geq 0,$$

where  $C(i) = p(\mathbf{X}(i), \mathbf{Y}(i), \psi_i\psi_i^*)$ , as  $C(i)^* X_s^a(i) C(i)$  is positive semidefinite since  $X_s^a(i)$  is positive semidefinite. In the same way one can check that  $L(p^* y_t^b p) \geq 0$  and  $L(p^* z p) \geq 0$ . That is, if we set

$$\mathcal{G} = \{x_s^a : s \in S, a \in A\} \cup \{y_t^b : t \in T, b \in B\} \cup \{z\},$$

then  $L$  is nonnegative (denoted as  $L \geq 0$ ) on the *truncated quadratic module*

$$\mathcal{M}_{2r}(\mathcal{G}) = \text{cone}\left\{p^*gp : p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle, g \in \mathcal{G} \cup \{1\}, \deg(p^*gp) \leq 2r\right\}. \quad (18)$$

Similarly, setting

$$\begin{aligned} \mathcal{H} = & \{z - z^2\} \cup \left\{1 - \sum_{a \in A} x_s^a : s \in S\right\} \\ & \cup \left\{1 - \sum_{b \in B} y_t^b : t \in T\right\} \cup \{[x_s^a, y_t^b] : (s, t, a, b) \in \Gamma\}, \end{aligned}$$

we have  $L = 0$  on the *truncated ideal*

$$\mathcal{I}_{2r}(\mathcal{H}) = \left\{ph : p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle, h \in \mathcal{H}, \deg(ph) \leq 2r\right\}. \quad (19)$$

Moreover, we have  $L(z) = \sum_i \lambda_i \text{Re}(\text{Tr}(\psi_i \psi_i^*)) = 1$ . In addition, for any matrices  $U, V \in \mathbb{C}^{d_i \times d_i}$  we have

$$\psi_i \psi_i^* U \psi_i \psi_i^* V \psi_i \psi_i^* = \psi_i \psi_i^* V \psi_i \psi_i^* U \psi_i \psi_i^*,$$

and therefore, in particular,

$$L(wzuzvz) = L(wzvvuz) \quad \text{for all } u, v, w \in \langle \mathbf{x}, \mathbf{y}, z \rangle \quad \text{with } \deg(wzuzvz) \leq 2r.$$

That is, we have  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{R}_r)$ , where

$$\mathcal{R}_r = \{zuzvz - zvzuz : u, v \in \langle \mathbf{x}, \mathbf{y}, z \rangle \text{ with } \deg(zuzvz) \leq 2r\}.$$

We get the idea of adding these last constraints from [37], where this is used to study the mutually unbiased bases problem.

We call  $\mathcal{M}(\mathcal{G}) = \mathcal{M}_\infty(\mathcal{G})$  the quadratic module generated by  $\mathcal{G}$ , and we call  $\mathcal{I}(\mathcal{H} \cup \mathcal{R}_\infty) = \mathcal{I}_\infty(\mathcal{H} \cup \mathcal{R}_\infty)$  the ideal generated by  $\mathcal{H} \cup \mathcal{R}_\infty$ .

For  $r \in \mathbb{N} \cup \{\infty\}$  we can now define the parameter:

$$\begin{aligned} \xi_r^q(P) = \min & \left\{L(1) : L \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_{2r}^* \text{ tracial and symmetric,} \right. \\ & L(z) = 1, L(x_s^a y_t^b z) = P(a, b|s, t) \text{ for all } (a, b, s, t) \in \Gamma, \\ & \left. L \geq 0 \text{ on } \mathcal{M}_{2r}(\mathcal{G}), L = 0 \text{ on } \mathcal{I}_{2r}(\mathcal{H} \cup \mathcal{R}_r)\right\}. \end{aligned}$$

Note that for order  $r = 1$  we get the trivial bound  $\xi_1^q(P) = 1$ .

For each finite  $r \in \mathbb{N}$  the parameter  $\xi_r^q(P)$  can be computed by semidefinite programming. Indeed, the condition  $L \geq 0$  on  $\mathcal{M}_{2r}(\mathcal{G})$  means that  $L(p^*gp) \geq 0$  for all  $g \in \mathcal{G} \cup \{1\}$  and all polynomials  $p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle$  with degree at most  $r - \lceil \deg(g)/2 \rceil$ . This is equivalent to requiring that the matrices  $(L(w^*gw'))$ , indexed by all words

$w, w'$  with degree at most  $r - \lceil \deg(g)/2 \rceil$ , are positive semidefinite. To see this, write  $p = \sum_w p_w w$  and let  $\hat{p} = (p_w)$  denote the vector of coefficients, then  $L(p^* g p) \geq 0$  is equivalent to  $\hat{p}^\top (L(w^* g w')) \hat{p} \geq 0$ . When  $g = 1$ , the matrix  $(L(w^* g w'))$  is indexed by the words of degree at most  $r$ , it is called the *moment matrix* of  $L$  and denoted by  $M_r(L)$  (or  $M(L)$  when  $r = \infty$ ). The entries of the matrices  $(L(w^* g w'))$  are linear combinations of the entries of  $M_r(L)$ , and the constraint  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{H} \cup \mathcal{K}_r)$  can be written as a set of linear constraints on the entries of  $M_r(L)$ . It follows that for finite  $r \in \mathbb{N}$ , the parameter  $\xi_r^q(P)$  is indeed computable by a semidefinite program.

Additionally, we define the parameter  $\xi_*^q(P)$  by adding to the definition of  $\xi_\infty^q(P)$  the constraint  $\text{rank}(M(L)) < \infty$ . By construction this gives a hierarchy of lower bounds for  $A_{qc}(P)$ :

$$\xi_1^q(P) \leq \dots \leq \xi_r^q(P) \leq \xi_\infty^q(P) \leq \xi_*^q(P) \leq A_{qc}(P).$$

Indeed, if  $L \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_{2r}^*$  is feasible for  $\xi_r^q(P)$  then its restriction to  $\mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_{2r-2}^*$  is feasible for  $\xi_{r-1}^q(P)$ , which implies  $\xi_{r-1}^q(P) \leq L(1)$  and thus  $\xi_{r-1}^q(P) \leq \xi_r^q(P)$ .

### 3.3 Background on positive tracial linear forms

Before we show the convergence results for the hierarchy  $\{\xi_r^q(P)\}$  we give some background on positive tracial linear forms, which we will use again in Sect. 4. We state these results using the variables  $x_1, \dots, x_n$ , where we use the notation  $\langle \mathbf{x} \rangle = \langle x_1, \dots, x_n \rangle$ . The results stated below do not always appear in this way in the sources cited; we follow the presentation of [16], where full proofs for all these results are also provided.

First we need a few more definitions. A polynomial  $p \in \mathbb{R}\langle \mathbf{x} \rangle$  is called symmetric if  $p^* = p$ , and we denote the set of symmetric polynomials by  $\text{Sym } \mathbb{R}\langle \mathbf{x} \rangle$ . Given  $\mathcal{G} \subseteq \text{Sym } \mathbb{R}\langle \mathbf{x} \rangle$  and  $\mathcal{H} \subseteq \mathbb{R}\langle \mathbf{x} \rangle$ , the set  $\mathcal{M}(\mathcal{G}) + \mathcal{I}(\mathcal{H})$  is called *Archimedean* if it contains the polynomial  $R - \sum_{i=1}^n x_i^2$  for some  $R > 0$ .

Recall that for our purposes a  $C^*$ -algebra  $\mathcal{A}$  can be defined as a norm closed  $*$ -subalgebra of the space  $\mathcal{B}(H)$  of bounded operators on a complex Hilbert space  $H$ . We say that a  $C^*$ -algebra  $\mathcal{A}$  is *unital* if it contains the identity operator (denoted 1). An element  $a \in \mathcal{A}$  is called *positive* if  $a = b^*b$  for some  $b \in \mathcal{A}$ . A linear form  $\tau$  on a unital  $C^*$ -algebra  $\mathcal{A}$  is said to be a *state* if  $\tau(1) = 1$  and  $\tau$  is positive; that is,  $\tau(a) \geq 0$  for all positive elements  $a \in \mathcal{A}$ . We say that a state  $\tau$  is *tracial* if  $\tau(ab) = \tau(ba)$  for all  $a, b \in \mathcal{A}$ . See, for example, [4] for more information on  $C^*$ -algebras.

The first result relates positive tracial linear forms to  $C^*$ -algebras; see [37] for the noncommutative (eigenvalue) setting and [8] for the tracial setting.

**Theorem 2** *Let  $\mathcal{G} \subseteq \text{Sym } \mathbb{R}\langle \mathbf{x} \rangle$  and  $\mathcal{H} \subseteq \mathbb{R}\langle \mathbf{x} \rangle$  and assume that  $\mathcal{M}(\mathcal{G}) + \mathcal{I}(\mathcal{H})$  is Archimedean. For a linear form  $L \in \mathbb{R}\langle \mathbf{x} \rangle^*$ , the following are equivalent:*

- (1)  *$L$  is symmetric, tracial, nonnegative on  $\mathcal{M}(\mathcal{G})$ , zero on  $\mathcal{I}(\mathcal{H})$ , and  $L(1) = 1$ ;*
- (2) *there is a unital  $C^*$ -algebra  $\mathcal{A}$  with tracial state  $\tau$  and  $\mathbf{X} \in \mathcal{A}^n$  such that  $g(\mathbf{X})$  is positive in  $\mathcal{A}$  for all  $g \in \mathcal{G}$ , and  $h(\mathbf{X}) = 0$  for all  $h \in \mathcal{H}$ , with*

$$L(p) = \tau(p(\mathbf{X})) \text{ for all } p \in \mathbb{R}\langle \mathbf{x} \rangle. \quad (20)$$

The following can be seen as the finite dimensional analogue of the above result. The proof of the unconstrained case ( $\mathcal{G} = \mathcal{H} = \emptyset$ ) can be found in [7], and for the constrained case in [8].

Given a linear form  $L \in \mathbb{R}\langle \mathbf{x} \rangle^*$ , recall that its moment matrix  $M(L)$  is given by  $M(L)_{u,v} = L(u^*v)$  for  $u, v \in \langle \mathbf{x} \rangle$ . Recall also that  $L$  is called a *normalized trace evaluation* if there exists a tuple  $(X_1, \dots, X_n)$  of  $d \times d$  Hermitian matrices (for some  $d \in \mathbb{N}$ ) such that  $L(p) = \text{Tr}(p(X_1, \dots, X_n))/d$  for all  $p \in \mathbb{R}\langle \mathbf{x} \rangle$ .

**Theorem 3** *Let  $\mathcal{G} \subseteq \text{Sym } \mathbb{R}\langle \mathbf{x} \rangle$  and  $\mathcal{H} \subseteq \mathbb{R}\langle \mathbf{x} \rangle$ . For  $L \in \mathbb{R}\langle \mathbf{x} \rangle^*$ , the following are equivalent:*

- (1)  *$L$  is a symmetric, tracial, linear form with  $L(1) = 1$  that is nonnegative on  $\mathcal{M}(\mathcal{G})$ , zero on  $\mathcal{I}(\mathcal{H})$ , and has  $\text{rank}(M(L)) < \infty$ ;*
- (2) *there is a finite dimensional  $C^*$ -algebra  $\mathcal{A}$  with a tracial state  $\tau$  and  $\mathbf{X} \in \mathcal{A}^n$  satisfying (20), with  $g(\mathbf{X})$  positive in  $\mathcal{A}$  for all  $g \in \mathcal{G}$  and  $h(\mathbf{X}) = 0$  for all  $h \in \mathcal{H}$ ;*
- (3)  *$L$  is a convex combination of normalized trace evaluations at tuples  $\mathbf{X}$  of Hermitian matrices that satisfy  $g(\mathbf{X}) \geq 0$  for all  $g \in \mathcal{G}$  and  $h(\mathbf{X}) = 0$  for all  $h \in \mathcal{H}$ .*

Given an integer  $r \in \mathbb{N}$  a (truncated) linear functional  $L \in \mathbb{R}\langle \mathbf{x} \rangle_{2r}$  is called  $\delta$ -flat if the principal submatrix  $M_{r-\delta}(L)$  of  $M_r(L)$  indexed by monomials up to degree  $r - \delta$  has the same rank as  $M_r(L)$ , i.e.,

$$\text{rank}(M_r(L)) = \text{rank}(M_{r-\delta}(L)). \quad (21)$$

One says  $L$  is *flat* if it is  $\delta$ -flat for some  $\delta \geq 1$ . The following result claims that any flat linear functional on a truncated polynomial space can be extended to a linear functional  $L$  on the full algebra of polynomials. It is due to Curto and Fialkow [11] in the commutative case and extensions to the noncommutative case can be found in [48] (for eigenvalue optimization) and [7, 22] (for trace optimization).

**Theorem 4** *Let  $1 \leq \delta \leq r < \infty$ ,  $\mathcal{G} \subseteq \text{Sym } \mathbb{R}\langle \mathbf{x} \rangle_{2\delta}$ , and  $\mathcal{H} \subseteq \mathbb{R}\langle \mathbf{x} \rangle_{2\delta}$ . If  $L \in \mathbb{R}\langle \mathbf{x} \rangle_{2r}^*$  is symmetric, tracial,  $\delta$ -flat, nonnegative on  $\mathcal{M}_{2r}(\mathcal{G})$ , and zero on  $\mathcal{I}_{2r}(\mathcal{H})$ , then  $L$  extends to a symmetric, tracial, linear form on  $\mathbb{R}\langle \mathbf{x} \rangle$  that is nonnegative on  $\mathcal{M}(\mathcal{G})$ , zero on  $\mathcal{I}(\mathcal{H})$ , and whose moment matrix  $M(L)$  has finite rank.*

The following technical lemma, based on the Banach-Alaoglu theorem, is a well-known tool to show asymptotic convergence results in polynomial optimization.

**Lemma 1** *Let  $\mathcal{G} \subseteq \text{Sym } \mathbb{R}\langle \mathbf{x} \rangle$ ,  $\mathcal{H} \subseteq \mathbb{R}\langle \mathbf{x} \rangle$ , and assume that for some  $d \in \mathbb{N}$  and  $R > 0$  we have  $R - (x_1^2 + \dots + x_n^2) \in \mathcal{M}_{2d}(\mathcal{G}) + \mathcal{I}_{2d}(\mathcal{H})$ . For  $r \in \mathbb{N}$  assume  $L_r \in \mathbb{R}\langle \mathbf{x} \rangle_{2r}^*$  is tracial, nonnegative on  $\mathcal{M}_{2r}(\mathcal{G})$  and zero on  $\mathcal{I}_{2r}(\mathcal{H})$ . Then  $|L_r(w)| \leq R^{|w|/2} L_r(1)$  for all  $w \in \langle \mathbf{x} \rangle_{2r-2d+2}$ . In addition, if  $\sup_r L_r(1) < \infty$ , then  $\{L_r\}_r$  has a pointwise converging subsequence in  $\mathbb{R}\langle \mathbf{x} \rangle^*$ .*

### 3.4 Convergence results

We first show that the parameter  $\xi_*^q(P)$  coincides with the average entanglement dimension  $A_q(P)$  and then we consider convergence properties of the bounds  $\xi_r^q(P)$  to the parameters  $\xi_\infty^q(P)$  and  $\xi_*^q(P)$ .

**Proposition 5** For any  $P \in C_q(\Gamma)$  we have  $\xi_*^q(P) = A_{qc}(P)$ .

*Proof* We already know  $\xi_*^q(P) \leq A_{qc}(P)$ . To show  $\xi_*^q(P) \geq A_{qc}(P)$  we let  $L$  be feasible for  $\xi_*^q(P)$ , so that  $L \geq 0$  on  $\mathcal{M}(\mathcal{G})$ ,  $L = 0$  on  $\mathcal{I}(\mathcal{H} \cup \mathcal{R}_\infty)$  and  $\text{rank}(M(L)) < \infty$ . We apply Theorem 3 to the scaled linear form  $L/L(1)$  (note that  $L(1) > 0$  since  $L(z) = 1$ ): there exist finitely many scalars  $\lambda_i \geq 0$  with  $\sum_i \lambda_i = L(1)$ , Hermitian matrix tuples  $\mathbf{X}(i) = (X_s^a(i))_{a,s}$  and  $\mathbf{Y}(i) = (Y_t^b(i))_{b,t}$ , and Hermitian matrices  $Z_i$ , so that

$$g(\mathbf{X}(i), \mathbf{Y}(i), Z_i) \geq 0 \text{ for all } g \in \mathcal{G}, \quad h(\mathbf{X}(i), \mathbf{Y}(i), Z_i) = 0 \text{ for all } h \in \mathcal{H} \cup \mathcal{R}_\infty, \quad (22)$$

and

$$L(p) = \sum_i \lambda_i \text{Tr}(p(\mathbf{X}(i), \mathbf{Y}(i), Z_i)) \text{ for all } p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle. \quad (23)$$

By Artin–Wedderburn theory (Theorem 1) we know that for each  $i$  there is a unitary matrix  $V_i$  such that  $V_i \mathbb{C}\langle \mathbf{X}(i), \mathbf{Y}(i), Z_i \rangle V_i^* = \bigoplus_k \mathbb{C}^{d_k \times d_k} \otimes I_{m_k}$ . Hence, after applying this further block diagonalization we may assume that in the decomposition (23), for each  $i$ ,  $\mathbb{C}\langle \mathbf{X}(i), \mathbf{Y}(i), Z_i \rangle$  is a full matrix algebra  $\mathbb{C}^{d_i \times d_i}$ .

Since  $h(\mathbf{X}(i), \mathbf{Y}(i), Z_i) = 0$  for all  $h \in R_\infty \cup \{z - z^2\}$ ,  $Z_i$  is a projector and the commutator  $[Z_i u Z_i, Z_i v Z_i]$  vanishes for all  $u, v \in \langle \mathbf{X}(i), \mathbf{Y}(i), Z_i \rangle$  and hence for all  $u, v \in \mathbb{C}\langle \mathbf{X}(i), \mathbf{Y}(i), Z_i \rangle$ . This means that  $[Z_i T_1 Z_i, Z_i T_2 Z_i] = 0$  for all  $T_1, T_2 \in \mathbb{C}^{d_i \times d_i}$ . As  $Z_i$  is a projector, there exists a unitary matrix  $U_i$  such that  $U_i Z_i U_i^* = \text{Diag}(1, \dots, 1, 0, \dots, 0)$ . The above then implies that for all  $T_1$  and  $T_2$ , the leading principal submatrices of size  $\text{rank}(Z_i)$  of  $U_i T_1 U_i^*$  and  $U_i T_2 U_i^*$  commute. This implies  $\text{rank}(Z_i) \leq 1$  and thus  $\text{Tr}(Z_i) \in \{0, 1\}$ . Let  $I$  be the set of indices with  $\text{Tr}(Z_i) = 1$ . Then we have  $\sum_{i \in I} \lambda_i = \sum_i \lambda_i \text{Tr}(Z_i) = L(z) = 1$ .

For each  $i \in I$  define  $P_i = (\text{Tr}(X_s^a(i) Y_t^b(i) Z_i))$ , which is a quantum correlation in  $C_{qc}^{d_i}(\Gamma)$  because  $\text{Tr}(Z_i) = 1$ , and  $X_s^a, Y_t^b \geq 0$  with  $\sum_a X_s^a(i) = \sum_b Y_t^b(i) = I$  and  $[X_s^a(i), Y_t^b(i)] = 0$  in view of (22). Using (23) we obtain  $P = \sum_{i \in I} \lambda_i P_i$ . Hence,  $(P_i, \lambda_i)_{i \in I}$  forms a feasible solution to  $A_{qc}(P)$  with objective value  $\sum_{i \in I} \lambda_i D_{qc}(P_i) \leq \sum_{i \in I} \lambda_i d_i \leq \sum_i \lambda_i d_i = L(1)$ .  $\square$

The problem  $\xi_r^q(P)$  differs in two ways from a standard tracial optimization problem. First it does not have the normalization  $L(1) = 1$  (and instead it minimizes  $L(1)$ ), and second it has ideal constraints  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{R}_r)$  where  $\mathcal{R}_r$  depends on the relaxation order  $r$ . Nevertheless we can show that asymptotic convergence still holds.

**Proposition 6** For any  $P \in C_q(\Gamma)$  we have  $\xi_r^q(P) \rightarrow \xi_\infty^q(P)$  as  $r \rightarrow \infty$ .

*Proof* First observe that  $1 - z^2, 1 - (x_s^a)^2, 1 - (y_t^b)^2 \in \mathcal{M}_4(\mathcal{G} \cup \mathcal{H}_0)$ , where  $\mathcal{H}_0$  contains the symmetric polynomials in  $\mathcal{H}$ ; i.e., omitting the commutators  $[x_s^a, y_t^b]$ . Indeed, we have  $1 - z^2 = (1 - z)^2 + 2(z - z^2)$  and

$$1 - (x_s^a)^2 = (1 - x_s^a)^2 + 2(1 - x_s^a)x_s^a(1 - x_s^a) + 2x_s^a \left( \left( 1 - \sum_{a'} x_s^{a'} \right) + \sum_{a' \neq a} x_s^{a'} \right) x_s^a,$$



and the same for  $y_t^b$ . Hence  $R - z^2 - \sum_{a,s} (x_s^a)^2 - \sum_{b,t} (y_t^b)^2 \in \mathcal{M}_4(\mathcal{G} \cup \mathcal{H}_0)$  for some  $R > 0$ . Fix  $\varepsilon > 0$  and for each  $r \in \mathbb{N}$  let  $L_r$  be feasible for  $\xi_r^q(P)$  with value  $L_r(1) \leq \xi_r^q(P) + \varepsilon$ . As  $L_r$  is tracial and zero on  $\mathcal{I}_{2r}(\mathcal{H}_0)$ , it follows (using the identity  $p^*gp = pp^*g + [p^*g, p]$ ) that  $L = 0$  on  $\mathcal{M}_{2r}(\mathcal{H}_0)$ . Hence,  $L_r \geq 0$  on  $\mathcal{M}_{2r}(\mathcal{G} \cup \mathcal{H}_0)$ . Since  $\sup_r L_r(1) \leq A_q(P) + \varepsilon$ , we can apply Lemma 1 and conclude that  $\{L_r\}_r$  has a converging subsequence; denote its limit by  $L_\varepsilon \in \mathbb{R}(\mathbf{x})^*$ . One can verify that  $L_\varepsilon$  is feasible for  $\xi_\infty^q(P)$ , and  $\xi_\infty^q(P) \leq L_\varepsilon(1) \leq \lim_{r \rightarrow \infty} \xi_r^q(P) + \varepsilon \leq \xi_\infty^q(P) + \varepsilon$ . Letting  $\varepsilon \rightarrow 0$  we obtain that  $\xi_\infty^q(P) = \lim_{r \rightarrow \infty} \xi_r^q(P)$ .  $\square$

Next we show that finite convergence holds under a certain flatness condition: if  $\xi_r^q(P)$  admits a  $\delta$ -flat optimal solution with  $\delta = \lceil r/3 \rceil + 1$ , then  $\xi_r^q(P) = \xi_*^q(P)$ . This result is a variation of the flat extension result from Theorem 4, where  $\delta$  now depends on the order  $r$  because the ideal constraints in  $\xi_r^q(P)$  depend on  $r$ .

**Proposition 7** *If  $\xi_r^q(P)$  admits a  $(\lceil r/3 \rceil + 1)$ -flat optimal solution, then we have  $\xi_r^q(P) = \xi_*^q(P)$ .*

*Proof* Let  $\delta = \lceil r/3 \rceil + 1$  and let  $L$  be a  $\delta$ -flat optimal solution to  $\xi_r^q(P)$ , i.e., such that  $\text{rank}(M_r(L)) = \text{rank}(M_{r-\delta}(L))$ . We have to show  $\xi_r^q(P) \geq \xi_*^q(P)$ , which we do by constructing a feasible solution  $\hat{L}$  to  $\xi_*^q(P)$  with the same objective value  $\hat{L}(1) = L(1)$ . In the proof of Theorem 4 (see [16, Thm. 2.3], and also [22, Prop. 6.1] for the original proof of this theorem), the linear form  $L$  is extended to a tracial symmetric linear form  $\hat{L}$  on  $\mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle$  that is nonnegative on  $\mathcal{M}(\mathcal{G})$ , zero on  $\mathcal{I}(\mathcal{H})$ , with  $\text{rank}(M(\hat{L})) < \infty$ . To do this a subset  $W$  of  $\langle \mathbf{x}, \mathbf{y}, z \rangle_{r-\delta}$  is found such that we have the vector space direct sum  $\mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle = \text{span}(W) \oplus \mathcal{I}(N_r(L))$ , where  $N_r(L)$  is the vector space

$$N_r(L) = \{p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_r : L(qp) = 0 \text{ for all } q \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle_r\}.$$

It is moreover shown that  $\mathcal{I}(N_r(L)) \subseteq N(\hat{L})$ . For  $p \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle$  we denote by  $r_p$  the unique element in  $\text{span}(W)$  such that  $p - r_p \in \mathcal{I}(N_r(L))$ .

We show that  $\hat{L}$  is zero on  $\mathcal{I}(\mathcal{R}_\infty)$ . Fix  $u, v, w \in \mathbb{R}\langle \mathbf{x}, \mathbf{y}, z \rangle$ . Then we have

$$\hat{L}(w(zuzvz - zvzuz)) = \hat{L}(wzuzvz) - \hat{L}(wzvzuz).$$

Since  $\hat{L}$  is tracial and  $u - r_u, v - r_v, w - r_w \in \mathcal{I}(N_r(L)) \subseteq N(\hat{L})$ , we have

$$\hat{L}(wzuzvz) = \hat{L}(r_w z r_u z r_v z) \quad \text{and} \quad \hat{L}(wzvzuz) = \hat{L}(r_w z r_v z r_u z).$$

Since  $\deg(r_u z r_v z r_w z) = \deg(r_v z r_u z r_w z) \leq 3 + 3(r - \delta) \leq 2r$  we have

$$\hat{L}(r_w z r_u z r_v z) = L(r_w z r_u z r_v z) \quad \text{and} \quad \hat{L}(r_w z r_v z r_u z) = L(r_w z r_v z r_u z).$$

So  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{R}_r)$  implies  $\hat{L} = 0$  on  $\mathcal{I}(\mathcal{R}_\infty)$ .

Since  $\hat{L}$  extends  $L$  we have  $\hat{L}(z) = L(z) = 1$  and  $\hat{L}(x_s^a y_t^b z) = L(x_s^a y_t^b z) = P(a, b|s, t)$  for all  $a, b, s, t$ . So,  $\hat{L}$  is feasible for  $\xi_*^q(P)$  and has the same objective value  $\hat{L}(1) = L(1)$ .  $\square$

## 4 Bounding quantum graph parameters

We investigate the quantum graph parameters  $\alpha_q(G)$ ,  $\gamma_q(G)$ ,  $\alpha_{qc}(G)$ , and  $\chi_{qc}(G)$ , which are quantum analogues of the classical graph parameters  $\alpha(G)$  and  $\chi(G)$ . They were introduced earlier in Sect. 1.3 in terms of nonlocal games and synchronous quantum correlations (in the tensor and commuting models). As we will see below, they can be reformulated in terms of the existence of positive semidefinite matrices with arbitrary size (or operators) satisfying a system of equations corresponding to the natural integer linear programming formulation of  $\alpha(G)$  and  $\chi(G)$ . This opens the way to using techniques from noncommutative polynomial optimization for designing hierarchies of bounds for the quantum graph parameters. We present these approaches and compare them with known hierarchies for the classical graph parameters.

### 4.1 Hierarchies $\gamma_r^{\text{col}}(G)$ and $\gamma_r^{\text{stab}}(G)$ based on synchronous correlations

In Sect. 1.3 we introduced quantum chromatic numbers (Definition 1) and quantum stability numbers (Definition 2) in terms of synchronous quantum correlations satisfying certain linear constraints. We first give (known) reformulations in terms of  $C^*$ -algebras, and then we reformulate those in terms of tracial optimization, which leads to the hierarchies  $\gamma_r^{\text{col}}(G)$  and  $\gamma_r^{\text{stab}}(G)$ .

The following result from [46] allows us to write a synchronous quantum correlation in terms of  $C^*$ -algebras admitting a tracial state.

**Theorem 5** ([46]) *Let  $\Gamma = A^2 \times S^2$  and  $P \in \mathbb{R}^\Gamma$ . We have  $P \in C_{qc,s}(\Gamma)$  (resp.,  $P \in C_{q,s}(\Gamma)$ ) if and only if there exists a unital (resp., finite dimensional)  $C^*$ -algebra  $\mathcal{A}$  with a faithful tracial state  $\tau$  and a set of projectors  $\{X_s^a : s \in S, a \in A\}$  in  $\mathcal{A}$  satisfying  $\sum_{a \in A} X_s^a = 1$  for all  $s \in S$  and  $P(a, b|s, t) = \tau(X_s^a X_t^b)$  for all  $s, t \in S$  and  $a, b \in A$ .*

Here we add the condition that  $\tau$  is faithful, that is,  $\tau(X^*X) = 0$  implies  $X = 0$ , since it follows from the GNS construction in the proof of [46]. This means that

$$0 = P(a, b|s, t) = \tau(X_s^a X_t^b) = \tau\left((X_s^a)^2 (X_t^b)^2\right) = \tau\left(\left(X_s^a X_t^b\right)^* X_s^a X_t^b\right)$$

implies  $X_s^a X_t^b = 0$ . It follows from Definition 1 and the above that  $\chi_{qc}(G)$  is equal to the smallest  $k \in \mathbb{N}$  for which there exists a  $C^*$ -algebra  $\mathcal{A}$ , a tracial state  $\tau$  on  $\mathcal{A}$ , and a family of projectors  $\{X_i^c : i \in V, c \in [k]\} \subseteq \mathcal{A}$  satisfying

$$\sum_{c \in [k]} X_i^c - 1 = 0 \quad \text{for all } i \in V, \quad (24)$$

$$X_i^c X_j^{c'} = 0 \quad \text{if } (c \neq c' \text{ and } i = j) \text{ or } (c = c' \text{ and } \{i, j\} \in E). \quad (25)$$

The quantum chromatic number  $\chi_q(G)$  is equal to the smallest  $k \in \mathbb{N}$  for which there exists a finite dimensional  $C^*$ -algebra  $\mathcal{A}$  with the above properties.

Analogously,  $\alpha_{qc}(G)$  is equal to the largest  $k \in \mathbb{N}$  for which there is a  $C^*$ -algebra  $\mathcal{A}$ , a tracial state  $\tau$  on  $\mathcal{A}$ , and a set of projectors  $\{X_c^i : c \in [k], i \in V\} \subseteq \mathcal{A}$  satisfying

$$\sum_{i \in V} X_c^i - 1 = 0 \quad \text{for all } c \in [k], \quad (26)$$

$$X_c^i X_{c'}^j = 0 \quad \text{if } (i \neq j \text{ and } c = c') \quad \text{or} \quad ((i = j \text{ or } \{i, j\} \in E) \text{ and } c \neq c'), \quad (27)$$

and  $\alpha_q(G)$  is equal to the largest  $k \in \mathbb{N}$  for which  $\mathcal{A}$  can be taken finite dimensional.

These reformulations of  $\chi_q(G)$ ,  $\chi_{qc}(G)$ ,  $\alpha_q(G)$  and  $\alpha_{qc}(G)$  also follow from [41, Thm. 4.7], where general quantum graph homomorphisms are considered; the formulations of  $\chi_q(G)$  and  $\chi_{qc}(G)$  are also made explicit in [41, Thm. 4.12].

*Remark 1* The above definition for the parameters  $\alpha_q(G)$  and  $\chi_q(G)$  (tensor model) can be simplified. Indeed, instead of asking for projectors  $\{X_i^c\}$  living in a finite dimensional  $C^*$ -algebra equipped with a tracial state and satisfying the constraints (24)–(25) or (26)–(27), one may ask for such projectors that are *matrices* of unspecified (but finite) size (as in [9, 32, 52]). This can be seen in the following two ways.

A first possibility is to apply Artin–Wedderburn theory, which tells us that any finite dimensional  $C^*$ -algebra is isomorphic to a matrix algebra.

An alternative, more elementary way is to use the link presented in Sect. 2 between synchronous quantum correlations and completely positive semidefinite matrices. Indeed, as we have seen there, having a synchronous quantum correlation  $P = (P(c, c'|i, j)) \in \mathbb{R}^{V^2 \times [k]^2}$  certifying  $\chi_q(G) \leq k$  is equivalent to having a set of positive semidefinite matrices  $\{X_i^c\}$  satisfying the constraints (24)–(25). Here we use the basic fact that since  $X_i^c, X_j^{c'} \succeq 0$ , we have  $P(c, c'|i, j) = \text{Tr}(X_i^c X_j^{c'}) = 0$  if and only if  $X_i^c X_j^{c'} = 0$ . Next, observe that the constraints (24)–(25) imply that the matrices  $X_i^c$  are projectors. Indeed, for every  $i, c'$ , by multiplying (24) by  $X_i^{c'}$  and using (25) we obtain  $(X_i^{c'})^2 = X_i^{c'}$ . The analogous result holds of course for the quantum stability number  $\alpha_q(G)$ .

Finally, note that restricting to scalar solutions ( $1 \times 1$  matrices) in these feasibility problems recovers the classical graph parameters  $\chi(G)$  and  $\alpha(G)$ .

We now reinterpret the above formulations in terms of tracial optimization. Given a graph  $G = (V, E)$ , let  $i \simeq j$  denote  $\{i, j\} \in E$  or  $i = j$ . For  $k \in \mathbb{N}$ , let  $\mathcal{H}_{G,k}^{\text{col}}$  and  $\mathcal{H}_{G,k}^{\text{stab}}$  denote the sets of polynomials corresponding to equations (24)–(25) and (26)–(27):

$$\begin{aligned} \mathcal{H}_{G,k}^{\text{col}} &= \left\{ 1 - \sum_{c \in [k]} x_i^c : i \in V \right\} \\ &\quad \cup \left\{ x_i^c x_j^{c'} : (c \neq c' \text{ and } i = j) \text{ or } (c = c' \text{ and } \{i, j\} \in E) \right\}, \\ \mathcal{H}_{G,k}^{\text{stab}} &= \left\{ 1 - \sum_{i \in V} x_c^i : c \in [k] \right\} \cup \left\{ x_c^i x_{c'}^j : (i \neq j \text{ and } c = c') \text{ or } (i \simeq j \text{ and } c \neq c') \right\}. \end{aligned}$$

We have

$$1 - (x_i^c)^2 \in \mathcal{M}_2(\emptyset) + \mathcal{I}_2\left(\mathcal{H}_{G,k}^{\text{col}}\right),$$

since  $1 - (x_i^c)^2 = (1 - x_i^c)^2 + 2(x_i^c - (x_i^c)^2)$ , and

$$x_i^c - (x_i^c)^2 = x_i^c \left(1 - \sum_{c'} x_i^{c'}\right) + \sum_{c': c' \neq c} x_i^c x_i^{c'} \in \mathcal{I}_2\left(\mathcal{H}_{G,k}^{\text{col}}\right), \quad (28)$$

and the analogous statements hold for  $\mathcal{H}_{G,k}^{\text{stab}}$ . Hence, both  $\mathcal{M}(\emptyset) + \mathcal{I}\left(\mathcal{H}_k^{\text{col}}\right)$  and  $\mathcal{M}(\emptyset) + \mathcal{I}\left(\mathcal{H}_k^{\text{stab}}\right)$  are Archimedean and we can apply Theorems 2 and 3 to express the quantum graph parameters in terms of positive tracial linear functionals. Namely,

$$\chi_{qc}(G) = \min \left\{ k \in \mathbb{N} : L \in \mathbb{R}\langle \{x_i^c : i \in V, c \in [k]\} \rangle^* \text{ symmetric, tracial, positive, } \right. \\ \left. L(1) = 1, L = 0 \text{ on } \mathcal{I}(\mathcal{H}_{G,k}^{\text{col}}) \right\},$$

and  $\chi_q(G)$  is obtained by adding the constraint  $\text{rank}(M(L)) < \infty$ . Likewise,

$$\alpha_{qc}(G) = \max \left\{ k \in \mathbb{N} : L \in \mathbb{R}\langle \{x_c^i : c \in [k], i \in V\} \rangle^* \text{ symmetric, tracial, positive, } \right. \\ \left. L(1) = 1, L = 0 \text{ on } \mathcal{I}(\mathcal{H}_{G,k}^{\text{stab}}) \right\},$$

and  $\alpha_q(G)$  is given by this program with the additional constraint  $\text{rank}(M(L)) < \infty$ .

Starting from these formulations it is natural to define a hierarchy  $\{\gamma_r^{\text{col}}(G)\}$  of lower bounds on  $\chi_{qc}(G)$  and a hierarchy  $\{\gamma_r^{\text{stab}}(G)\}$  of upper bounds on  $\alpha_{qc}(G)$ , where the bounds of order  $r \in \mathbb{N}$  are obtained by truncating  $L$  to polynomials of degree at most  $2r$  and truncating the ideal to degree  $2r$ :

$$\gamma_r^{\text{col}}(G) = \min \left\{ k \in \mathbb{N} : L \in \mathbb{R}\langle \{x_i^c : i \in V, c \in [k]\} \rangle_{2r}^* \text{ symmetric, tracial, positive, } \right. \\ \left. L(1) = 1, L = 0 \text{ on } \mathcal{I}_{2r}\left(\mathcal{H}_{G,k}^{\text{col}}\right) \right\}, \\ \gamma_r^{\text{stab}}(G) = \max \left\{ k \in \mathbb{N} : L \in \mathbb{R}\langle \{x_c^i : c \in [k], i \in V\} \rangle_{2r}^* \text{ symmetric, tracial, positive, } \right. \\ \left. L(1) = 1, L = 0 \text{ on } \mathcal{I}_{2r}(\mathcal{H}_{G,k}^{\text{stab}}) \right\}.$$

Then, by defining  $\gamma_*^{\text{col}}(G)$  and  $\gamma_*^{\text{stab}}(G)$  by adding the constraint  $\text{rank}(M(L)) < \infty$  to  $\gamma_{\infty}^{\text{col}}(G)$  and  $\gamma_{\infty}^{\text{stab}}(G)$ , we have

$$\gamma_{\infty}^{\text{col}}(G) = \chi_{qc}(G), \quad \gamma_{\infty}^{\text{stab}}(G) = \alpha_{qc}(G), \text{ and} \\ \gamma_*^{\text{col}}(G) = \chi_q(G), \quad \gamma_*^{\text{stab}}(G) = \alpha_q(G).$$

The optimization problems  $\gamma_r^{\text{col}}(G)$ , for  $r \in \mathbb{N}$ , can be computed by semidefinite programming and binary search on  $k$ , since the positivity condition on  $L$  can

be expressed by requiring that its truncated moment matrix  $M_r(L) = (L(w^*w'))$  (indexed by words with degree at most  $r$ ) is positive semidefinite. If there is an optimal solution  $(k, L)$  to  $\gamma_r^{\text{col}}(G)$  with  $L$  flat, then, by Theorem 4, we have equality  $\gamma_r^{\text{col}}(G) = \chi_q(G)$ . Since  $\{\gamma_r^{\text{col}}(G)\}_{r \in \mathbb{N}}$  is a monotone nondecreasing sequence of lower bounds on  $\chi_q(G)$ , there exists an  $r_0$  such that for all  $r \geq r_0$  we have  $\gamma_r^{\text{col}}(G) = \gamma_{r_0}^{\text{col}}(G)$ , which is equal to  $\gamma_{\infty}^{\text{col}}(G) = \chi_{qc}(G)$  by Lemma 1. The analogous statements hold for the parameters  $\gamma_r^{\text{stab}}(G)$ . Hence, we have shown the following result.

**Proposition 8** *There is an  $r_0 \in \mathbb{N}$  such that  $\gamma_r^{\text{col}}(G) = \chi_{qc}(G)$  and  $\gamma_r^{\text{stab}}(G) = \alpha_{qc}(G)$  for all  $r \geq r_0$ . Moreover, if  $\gamma_r^{\text{col}}(G)$  admits a flat optimal solution, then  $\gamma_r^{\text{col}}(G) = \chi_q(G)$ , and if  $\gamma_r^{\text{stab}}(G)$  admits a flat optimal solution, then  $\gamma_r^{\text{stab}}(G) = \alpha_q(G)$ .*

**Remark 2** A hierarchy  $\{\mathcal{Q}_r(\Gamma)\}$  of semidefinite outer approximations for the set  $C_{qc}(\Gamma)$  of commuting quantum correlations was constructed in [46] (revisiting the approach in [36, 48]). This hierarchy converges, that is,

$$C_{qc}(\Gamma) = \mathcal{Q}_{\infty}(\Gamma) = \bigcap_{r \in \mathbb{N}} \mathcal{Q}_r(\Gamma).$$

These approximations  $\mathcal{Q}_r(\Gamma)$  are based on the eigenvalue optimization approach, applied to the formulation (4) of commuting quantum correlations. So they use linear functionals on polynomials involving the two sets of variables  $x_s^a$  and  $y_t^b$  for  $(a, b, s, t) \in \Gamma$ . Paulsen et al. [46] use these outer approximations to define a hierarchy of lower bounds converging to  $\chi_{qc}(G)$ , where the bounds are defined in terms of feasibility problems over the sets  $\mathcal{Q}_r(\Gamma)$ .

For synchronous correlations we can use the result of Theorem 5 and the tracial optimization approach used here to define directly a converging hierarchy  $\{\mathcal{Q}_{r,s}(\Gamma)\}$  of outer semidefinite approximations for the set  $C_{qc,s}(\Gamma)$  of synchronous commuting quantum correlations. These approximations now use linear functionals on polynomials involving only one set of variables  $x_s^a$  for  $(a, s) \in A \times S$ . Namely, for  $r \in \mathbb{N} \cup \{\infty\}$  define  $\mathcal{Q}_{r,s}(\Gamma)$  as the set of  $P \in \mathbb{R}^F$  for which there exists a symmetric, tracial, positive linear functional  $L \in \mathbb{R}\langle\{x_s^a : (a, s) \in A \times S\}\rangle_{2r}^*$  such that  $L(1) = 1$  and  $L = 0$  on the ideal generated by the polynomials  $x_s^a - (x_s^a)^2$  ( $(a, s) \in A \times S$ ) and  $1 - \sum_{a \in A} x_s^a$  ( $s \in S$ ), truncated at degree  $2r$ . Then we have

$$C_{qc,s}(\Gamma) = \mathcal{Q}_{\infty,s}(\Gamma) = \bigcap_{r \in \mathbb{N}} \mathcal{Q}_{r,s}(\Gamma).$$

The *synchronous value* of a nonlocal game is defined in [13] as the maximum value of the objective function (11) over the set  $C_{qc,s}(\Gamma)$ . By maximizing the objective (11) over the relaxations  $\mathcal{Q}_{r,s}(\Gamma)$  we get a hierarchy of semidefinite programming upper bounds that converges to the synchronous value of the game. Finally note that one can also view the parameters  $\gamma_r^{\text{col}}(G)$  as solving feasibility problems over the sets  $\mathcal{Q}_{r,s}(\Gamma)$ .

## 4.2 Hierarchies $\xi_r^{\text{col}}(G)$ and $\xi_r^{\text{stab}}(G)$ based on Lasserre type bounds

Here we revisit some known Lasserre type hierarchies for the classical stability number  $\alpha(G)$  and chromatic number  $\chi(G)$  and we show that their tracial noncommutative analogues can be used to recover known parameters such as the projective packing number  $\alpha_p(G)$ , the projective rank  $\xi_f(G)$ , and the tracial rank  $\xi_{\text{tr}}(G)$ . Compared to the hierarchies defined in the previous section, these Lasserre type hierarchies use less variables (they only use variables indexed by the vertices of the graph  $G$ ), but they also do not converge to the (commuting) quantum chromatic or stability number.

Given a graph  $G = (V, E)$ , define the set of polynomials

$$\mathcal{H}_G = \{x_i - x_i^2 : i \in V\} \cup \{x_i x_j : \{i, j\} \in E\}$$

in the variables  $\mathbf{x} = (x_i : i \in V)$  (which are commutative or noncommutative depending on the context). Note that  $1 - x_i^2 \in \mathcal{M}_2(\emptyset) + \mathcal{I}_2(\mathcal{H}_G)$  for all  $i \in V$ , so that  $\mathcal{M}(\emptyset) + \mathcal{I}(\mathcal{H}_G)$  is Archimedean.

### 4.2.1 Semidefinite programming bounds on the projective packing number

We first recall the Lasserre hierarchy of bounds for the classical stability number  $\alpha(G)$ . Starting from the formulation of  $\alpha(G)$  via the optimization problem

$$\alpha(G) = \sup \left\{ \sum_{i \in V} x_i : x \in \mathbb{R}^n, h(x) = 0 \text{ for } h \in \mathcal{H}_G \right\}, \quad (29)$$

the  $r$ -th level of the Lasserre hierarchy for  $\alpha(G)$  (introduced in [24, 26]) is defined by

$$\text{las}_r^{\text{stab}}(G) = \sup \left\{ L \left( \sum_{i \in V} x_i \right) : L \in \mathbb{R}[\mathbf{x}]_{2r}^* \text{ positive, } L(1) = 1, L = 0 \text{ on } \mathcal{I}_{2r}(\mathcal{H}_G) \right\}.$$

Then we have  $\text{las}_{r+1}^{\text{stab}}(G) \leq \text{las}_r^{\text{stab}}(G)$  and the first bound is Lovász' theta number:  $\text{las}_1^{\text{stab}}(G) = \vartheta(G)$ . Finite convergence to  $\alpha(G)$  is shown in [26]:

$$\text{las}_{\alpha(G)}^{\text{stab}}(G) = \alpha(G).$$

Roberson [51] introduces the *projective packing number*

$$\begin{aligned} \alpha_p(G) &= \sup \left\{ \frac{1}{d} \sum_{i \in V} \text{rank } X_i : d \in \mathbb{N}, \mathbf{X} \in (\mathcal{S}^d)^n \text{ projectors,} \right. \\ &\quad \left. X_i X_j = 0 \text{ for } \{i, j\} \in E \right\} \\ &= \sup \left\{ \frac{1}{d} \text{Tr} \left( \sum_{i \in V} X_i \right) : d \in \mathbb{N}, \mathbf{X} \in (\mathcal{S}^d)^n, h(\mathbf{X}) = 0 \text{ for } h \in \mathcal{H}_G \right\} \quad (30) \end{aligned}$$

as an upper bound for the quantum stability number  $\alpha_q(G)$ . Here  $\mathcal{S}^d$  denotes the set of real symmetric  $d \times d$  matrices. Note that the inequality  $\alpha_q(G) \leq \alpha_p(G)$  also follows from Proposition 9 below. Comparing (29) and (30) we see that the parameter  $\alpha_p(G)$  can be viewed as a noncommutative analogue of  $\alpha(G)$ .

For  $r \in \mathbb{N} \cup \{\infty\}$  we define the noncommutative analogue of  $\text{las}_r^{\text{stab}}(G)$  by

$$\xi_r^{\text{stab}}(G) = \sup \left\{ L \left( \sum_{i \in V} x_i \right) : L \in \mathbb{R} \langle \mathbf{x} \rangle_{2r}^* \text{ tracial, symmetric, and positive,} \right. \\ \left. L(1) = 1, L = 0 \text{ on } \mathcal{I}_{2r}(\mathcal{H}_G) \right\},$$

and  $\xi_*^{\text{stab}}(G)$  by adding the constraint  $\text{rank}(M(L)) < \infty$  to the definition of  $\xi_\infty^{\text{stab}}(G)$ .

In view of Theorems 2 and 3, both  $\xi_\infty^{\text{stab}}(G)$  and  $\xi_*^{\text{stab}}(G)$  can be reformulated in terms of  $C^*$ -algebras:  $\xi_\infty^{\text{stab}}(G)$  (resp.,  $\xi_*^{\text{stab}}(G)$ ) is the largest value of  $\tau(\sum_{i \in V} X_i)$ , where  $\mathcal{A}$  is a (resp., finite-dimensional)  $C^*$ -algebra with tracial state  $\tau$  and  $X_i \in \mathcal{A}$  ( $i \in [n]$ ) are projectors satisfying  $X_i X_j = 0$  for all  $\{i, j\} \in E$ . Moreover, as we now see, the parameter  $\xi_*^{\text{stab}}(G)$  coincides with the projective packing number and the parameters  $\xi_*^{\text{stab}}(G)$  and  $\xi_\infty^{\text{stab}}(G)$  upper bound the quantum stability numbers.

**Proposition 9** *We have  $\xi_*^{\text{stab}}(G) = \alpha_p(G) \geq \alpha_q(G)$  and  $\xi_\infty^{\text{stab}}(G) \geq \alpha_{qc}(G)$ .*

*Proof* By (30),  $\alpha_p(G)$  is the largest value of  $L(\sum_{i \in V} x_i)$  taken over all linear functionals  $L$  that are normalized trace evaluations at projectors  $\mathbf{X} \in (\mathcal{S}^d)^n$  (for some  $d \in \mathbb{N}$ ) with  $X_i X_j = 0$  for  $\{i, j\} \in E$ . By convexity the optimum remains unchanged when considering a convex combination of such trace evaluations. In view of Theorem 3 [the equivalence between (1) and (3)], we can conclude that this optimum value is precisely the parameter  $\xi_*^{\text{stab}}(G)$ . This shows equality  $\alpha_p(G) = \xi_*^{\text{stab}}(G)$ .

Consider a  $C^*$ -algebra  $\mathcal{A}$  with tracial state  $\tau$  and a set of projectors  $X_c^i \in \mathcal{A}$  (for  $i \in V$ ,  $c \in [k]$ ) satisfying (26)–(27). Then, setting  $X_i = \sum_{c \in [k]} X_c^i$  for  $i \in V$ , we obtain projectors  $X_i \in \mathcal{A}$  that satisfy  $X_i X_j = 0$  if  $\{i, j\} \in E$ . Moreover, the following holds:  $\tau(\sum_{i \in V} X_i) = \sum_{c \in [k]} \tau(\sum_{i \in V} X_c^i) = k$ . This shows  $\xi_\infty^{\text{stab}}(G) \geq \alpha_{qc}(G)$  and, when restricting  $\mathcal{A}$  to be finite dimensional,  $\xi_*^{\text{stab}}(G) \geq \alpha_q(G)$ .  $\square$

Using Lemma 1 one can verify that  $\xi_r^{\text{stab}}(G)$  converges to  $\xi_\infty^{\text{stab}}(G)$  as  $r \rightarrow \infty$ , and for  $r \in \mathbb{N} \cup \{\infty\}$  the infimum in  $\xi_r^{\text{stab}}(G)$  is attained. Moreover, by Theorem 4, if  $\xi_r^{\text{stab}}(G)$  admits a flat optimal solution, then equality  $\xi_r^{\text{stab}} = \xi_*^{\text{stab}}(G)$  holds. The first bound  $\xi_1^{\text{stab}}(G)$  coincides with the theta number, since  $\xi_1^{\text{stab}}(G) = \text{las}_1^{\text{stab}}(G) = \vartheta(G)$ . Summarizing we have  $\alpha_{qc}(G) \leq \xi_\infty^{\text{stab}}(G)$  and the following chain of inequalities

$$\alpha_q(G) \leq \alpha_p(G) = \xi_*^{\text{stab}}(G) \leq \xi_\infty^{\text{stab}}(G) \leq \xi_r^{\text{stab}}(G) \leq \xi_1^{\text{stab}}(G) = \vartheta(G).$$

#### 4.2.2 Semidefinite programming bounds on the projective rank and tracial rank

We now turn to the (quantum) chromatic numbers. First recall the definition of the fractional chromatic number:

$$\chi_f(G) := \min \left\{ \sum_{S \in \mathcal{S}_G} \lambda_S : \lambda \in \mathbb{R}_+^{\mathcal{S}_G}, \sum_{S \in \mathcal{S}_G: i \in S} \lambda_S = 1 \text{ for all } i \in V \right\}.$$



where  $\mathcal{S}_G$  is the set of stable sets of  $G$ . Clearly,  $\chi_f(G) \leq \chi(G)$ . The following Lasserre type lower bounds for the classical chromatic number  $\chi(G)$  are defined in [19]:

$$\text{las}_r^{\text{col}}(G) = \inf \left\{ L(1) : L \in \mathbb{R}[\mathbf{x}]_{2r}^* \text{ positive, } L(x_i) = 1 \ (i \in V), \right. \\ \left. L = 0 \text{ on } \mathcal{S}_{2r}(\mathcal{H}_G) \right\}.$$

Note that we may view  $\chi_f(G)$  as minimizing  $L(1)$  over all linear functionals  $L \in \mathbb{R}[\mathbf{x}]^*$  that are conic combinations of evaluations at characteristic vectors of stable sets. From this we see that  $\text{las}_r^{\text{col}}(G) \leq \chi_f(G)$  for all  $r \geq 1$ . In [19] it is shown that finite convergence to  $\chi_f(G)$  holds:

$$\text{las}_{\alpha(G)}^{\text{col}}(G) = \chi_f(G).$$

The bound of order  $r = 1$  coincides with the theta number:  $\text{las}_1^{\text{col}}(G) = \vartheta(\overline{G})$ .

The following parameter  $\xi_f(G)$ , called the *projective rank* of  $G$ , was introduced in [32] as a lower bound on the quantum chromatic number  $\chi_q(G)$ :

$$\xi_f(G) := \inf \left\{ \frac{d}{r} : d, r \in \mathbb{N}, X_1, \dots, X_n \in \mathcal{S}^d, \text{Tr}(X_i) = r \ (i \in V), \right. \\ \left. X_i^2 = X_i \ (i \in V), X_i X_j = 0 \ (\{i, j\} \in E) \right\}.$$

**Proposition 10** ([32]) *For any graph  $G$  we have  $\xi_f(G) \leq \chi_q(G)$ .*

*Proof* Set  $k = \chi_q(G)$ . It is shown in [9] that in the definition of  $\chi_q(G)$  from (24)–(25), one may assume w.l.o.g. that  $X_i^c$  are projectors that all have the same rank, say,  $r$ . Then, for any given color  $c \in [k]$ , the matrices  $X_i^c$  ( $i \in V$ ) provide a feasible solution to  $\xi_f(G)$  with value  $d/r$ . This shows  $\xi_f(G) \leq d/r$ . Finally,  $d/r = k$  holds since by (24)–(25) we have  $d = \text{rank}(I) = \sum_{c=1}^k \text{rank}(X_i^c) = kr$ .  $\square$

In [46, Prop. 5.11] it is shown that the projective rank can equivalently be defined as

$$\xi_f(G) = \inf \left\{ \lambda : \mathcal{A} \text{ is a finite dimensional } C^* \text{-algebra with tracial state } \tau, \right. \\ \left. X_i \in \mathcal{A} \text{ projector with } \tau(X_i) = 1/\lambda \ (i \in V), \right. \\ \left. X_i X_j = 0 \ (\{i, j\} \in E) \right\}.$$

Paulsen et al. [46] also define the *tracial rank*  $\xi_{tr}(G)$  of  $G$  as the parameter obtained by omitting in the above definition of  $\xi_f(G)$  the restriction that  $\mathcal{A}$  has to be finite dimensional. The motivation for the parameter  $\xi_{tr}(G)$  is that it lower bounds the *commuting* quantum chromatic number [46, Thm. 5.11]:

$$\xi_{tr}(G) \leq \chi_{qc}(G).$$

Using Theorems 2 and 3 (which we apply to  $L/L(1)$  when  $L$  is not normalized), we obtain the following reformulations:

$$\xi_f(G) = \inf \left\{ L(1) : L \in \mathbb{R}\langle \mathbf{x} \rangle^* \text{ tracial, symmetric, positive, } \text{rank}(M(L)) < \infty, \right. \\ \left. L(x_i) = 1 \ (i \in V), \ L = 0 \text{ on } \mathcal{I}(\mathcal{H}_G) \right\},$$

and  $\xi_{tr}(G)$  is obtained by the same program without the restriction  $\text{rank}(M(L)) < \infty$ . In addition, we obtain that in this formulation of  $\xi_f(G)$  we can equivalently optimize over all  $L$  that are conic combinations of trace evaluations at projectors  $X_i \in \mathbb{S}^d$  (for some  $d \in \mathbb{N}$ ) satisfying  $X_i X_j = 0$  for all  $\{i, j\} \in E$ . If we restrict the optimization to conic combinations of *scalar* evaluations ( $d = 1$ ) we obtain the fractional chromatic number. This shows that the projective rank can be seen as the noncommutative analogue of the fractional chromatic number, as was already observed in [32, 46].

The above formulations of the parameters  $\xi_{tr}(G)$  and  $\xi_f(G)$  in terms of linear functionals also show that they fit within the following hierarchy  $\{\xi_r^{\text{col}}(G)\}_{r \in \mathbb{N} \cup \{\infty\}}$ , defined as the noncommutative tracial analogue of the hierarchy  $\{\text{las}_r^{\text{col}}(G)\}_r$ :

$$\xi_r^{\text{col}}(G) = \inf \left\{ L(1) : L \in \mathbb{R}\langle \mathbf{x} \rangle_{2r}^* \text{ tracial, symmetric, and positive, } \right. \\ \left. L(x_i) = 1 \ (i \in V), \ L = 0 \text{ on } \mathcal{I}_{2r}(\mathcal{H}_G) \right\}.$$

Again,  $\xi_*^{\text{col}}(G)$  is the parameter obtained by adding the constraint  $\text{rank}(M(L)) < \infty$  to the program defining  $\xi_\infty^{\text{col}}(G)$ . By the above discussion the following holds.

**Proposition 11** *We have  $\xi_*^{\text{col}}(G) = \xi_f(G) \leq \chi_q(G)$  and  $\xi_\infty^{\text{col}}(G) = \xi_{tr}(G) \leq \chi_{qc}(G)$ .*

Using Lemma 1 one can verify that the parameters  $\xi_r^{\text{col}}(G)$  converge to  $\xi_\infty^{\text{col}}(G)$ . Moreover, by Theorem 4, if  $\xi_r^{\text{col}}(G)$  admits a flat optimal solution, then we have  $\xi_r^{\text{col}} = \xi_*^{\text{col}}(G)$ . Also, the parameter  $\xi_1^{\text{col}}(G)$  coincides with  $\text{las}_1^{\text{col}}(G) = \vartheta(\overline{G})$ . Summarizing we have  $\xi_\infty^{\text{col}}(G) = \xi_{tr}(G) \leq \chi_{qc}(G)$  and the following chain of inequalities

$$\vartheta(\overline{G}) = \xi_1^{\text{col}}(G) \leq \xi_r^{\text{col}}(G) \leq \xi_\infty^{\text{col}}(G) = \xi_{tr}(G) \leq \xi_*^{\text{col}}(G) = \xi_f(G) \leq \chi_q(G).$$

Observe that the bounds  $\text{las}_r^{\text{col}}(G)$  and  $\xi_r^{\text{col}}(G)$  remain below the fractional chromatic number  $\chi_f(G)$ , since  $\xi_f(G) = \xi_*^{\text{col}}(G) \leq \text{las}_*^{\text{col}}(G) = \chi_f(G)$ . Hence, these bounds are weak if  $\chi_f(G)$  is close to  $\vartheta(\overline{G})$  and far from  $\chi(G)$  or  $\chi_q(G)$ . In the classical setting this is the case, e.g., for the class of Kneser graphs  $G = K(n, r)$ , with vertex set the set of all  $r$ -subsets of  $[n]$  and having an edge between any two disjoint  $r$ -subsets. By results of Lovász [29, 30], the fractional chromatic number is  $n/r$ , which is known to be equal to  $\vartheta(\overline{K(n, r)})$ , while the chromatic number is  $n - 2r + 2$ . In [19] this was used as a motivation to define a new hierarchy of lower bounds  $\{\Lambda_r(G)\}$  on the chromatic number that can go beyond the fractional chromatic number. In Sect. 4.3 we recall this approach and show that its extension to the tracial setting recovers the hierarchy  $\{\gamma_r^{\text{col}}(G)\}$  introduced in Sect. 4.1. We also show how a similar technique can be used to recover the hierarchy  $\{\gamma_r^{\text{stab}}(G)\}$ .

#### 4.2.3 A link between $\xi_r^{\text{stab}}(G)$ and $\xi_r^{\text{col}}(G)$

In [19, Thm. 3.1] it is shown that the bounds  $\text{las}_r^{\text{stab}}(G)$  and  $\text{las}_r^{\text{col}}(G)$  satisfy

$$\text{las}_r^{\text{stab}}(G)\text{las}_r^{\text{col}}(G) \geq |V| \quad \text{for any } r \geq 1,$$

with equality if  $G$  is vertex-transitive. This extends a well-known property of the theta number (i.e., the case  $r = 1$ ). The same property holds for the noncommutative analogues  $\xi_r^{\text{stab}}(G)$  and  $\xi_r^{\text{col}}(G)$ .

**Lemma 2** *For a graph  $G = (V, E)$  and  $r \in \mathbb{N} \cup \{\infty, *\}$  we have  $\xi_r^{\text{stab}}(G)\xi_r^{\text{col}}(G) \geq |V|$ , with equality if  $G$  is vertex-transitive.*

*Proof* Let  $L$  be feasible for  $\xi_r^{\text{col}}(G)$ . Then  $\tilde{L} = L/L(1)$  provides a solution to  $\xi_r^{\text{stab}}(G)$  with value  $\tilde{L}(\sum_{i \in V} x_i) = |V|/L(1)$ , implying that  $\xi_r^{\text{stab}}(G) \geq |V|/L(1)$  and therefore  $\xi_r^{\text{stab}}(G)\xi_r^{\text{col}}(G) \geq |V|$ .

Assume  $G$  is vertex-transitive. Let  $L$  be a feasible solution for  $\xi_r^{\text{stab}}(G)$ . As  $G$  is vertex-transitive we may assume (after symmetrization) that  $L(x_i)$  takes a constant value. Set  $L(x_i) =: 1/\lambda$  for all  $i \in V$ , so that the objective value of  $L$  for  $\xi_r^{\text{stab}}(G)$  is  $|V|/\lambda$ . Then  $\tilde{L} = \lambda L$  provides a feasible solution for  $\xi_r^{\text{col}}(G)$  with value  $\lambda$ , implying  $\xi_r^{\text{col}}(G) \leq \lambda$ . This shows  $\xi_r^{\text{col}}(G)\xi_r^{\text{stab}}(G) \leq |V|$ .  $\square$

For a vertex-transitive graph  $G$ , the inequality  $\xi_f(G)\alpha_q(G) \leq |V|$  is shown in [32, Lem. 6.5]; it can be recovered from the  $r = *$  case of Lemma 2 and  $\alpha_q(G) \leq \alpha_p(G)$ .

#### 4.2.4 Comparison to existing semidefinite programming bounds

By adding the constraints  $L(x_i x_j) \geq 0$ , for all  $i, j \in V$ , to the program defining  $\xi_1^{\text{col}}(G)$ , we obtain the strengthened theta number  $\vartheta^+(\overline{G})$  (from [56]). Moreover, if we add the constraints

$$L(x_i x_j) \geq 0 \quad \text{for } i \neq j \in V, \quad (31)$$

$$\sum_{j \in C} L(x_i x_j) \leq 1 \quad \text{for } i \in V, \quad (32)$$

$$L(1) + \sum_{i \in C, j \in C'} L(x_i x_j) \geq |C| + |C'| \quad \text{for } C, C' \text{ distinct cliques in } G \quad (33)$$

to the program defining the parameter  $\xi_1^{\text{col}}(G)$ , then we obtain the parameter  $\xi_{\text{SDP}}(G)$ , which is introduced in [46, Thm. 7.3] as a lower bound on  $\xi_{\text{tr}}(G)$ . We will now show that the inequalities (31)–(33) are in fact valid for  $\xi_2^{\text{col}}(G)$ , which implies

$$\xi_2^{\text{col}}(G) \geq \xi_{\text{SDP}}(G) \geq \vartheta^+(\overline{G}).$$

For this, given a clique  $C$  in  $G$ , we define the polynomial

$$g_C := 1 - \sum_{i \in C} x_i \in \mathbb{R}\langle \mathbf{x} \rangle.$$

Then (32) and (33) can be reformulated as  $L(x_i g_C) \geq 0$  and  $L(g_C g_{C'}) \geq 0$ , respectively, using the fact that  $L(x_i) = L(x_i^2) = 1$  for all  $i \in V$ . Hence, to show that any feasible  $L$  for  $\xi_2^{\text{col}}(G)$  satisfies (31)–(33), it suffices to show Lemma 3 below. Recall that a commutator is a polynomial of the form  $[p, q] = pq - qp$  with  $p, q \in \mathbb{R}\langle \mathbf{x} \rangle$ . We denote by  $\Theta_r$  the set of linear combinations of commutators  $[p, q]$  with  $\deg(pq) \leq r$ .

**Lemma 3** *Let  $C$  and  $C'$  be cliques in a graph  $G$  and let  $i, j \in V$ . Then we have*

$$g_C \in \mathcal{M}_2(\emptyset) + \mathcal{I}_2(\mathcal{H}_G), \text{ and } x_i x_j, x_i g_C, g_C g_{C'} \in \mathcal{M}_4(\emptyset) + \mathcal{I}_4(\mathcal{H}_G) + \Theta_4.$$

*Proof* The claim  $g_C \in \mathcal{M}_2(\emptyset) + \mathcal{I}_2(\mathcal{H}_G)$  follows from the identity

$$g_C = \underbrace{\left(1 - \sum_{i \in C} x_i\right)^2}_{g_C} + \underbrace{\sum_{i \in C} (x_i - x_i^2) + \sum_{i \neq j \in C} x_i x_j}_h = g_C^2 + h, \quad (34)$$

where  $h \in \mathcal{I}_2(\mathcal{H}_G)$ . We also have

$$\begin{aligned} x_i x_j &= x_i x_j^2 x_i + x_j (x_i - x_i^2) + x_i^2 (x_j - x_j^2) + [x_i, x_i x_j^2] + [x_i - x_i^2, x_j], \\ x_i g_C &= x_i g_C^2 x_i + g_C^2 (x_i - x_i^2) + [x_i - x_i^2, g_C^2] + [x_i, x_i g_C^2], \end{aligned}$$

and, writing analogously  $g_{C'} = g_{C'}^2 + h'$  with  $h' \in \mathcal{I}_2(\mathcal{H}_G)$ , we have

$$g_C g_{C'} = g_C g_{C'}^2 g_C + [g_C, g_C g_{C'}^2] + [h, g_{C'}^2] + g_C^2 h' + h h' + g_C^2 h. \quad \square$$

Using the bound  $\xi_{\text{SDP}}(G)$  it is shown in [46, Thm. 7.4] that the tracial rank of the cycle  $C_{2n+1}$  satisfies  $\xi_{\infty}^{\text{col}}(C_{2n+1}) = (2n+1)/n$ . Combining this with Lemma 2 gives  $n = \xi_{\infty}^{\text{stab}}(C_{2n+1}) \geq \alpha_{qc}(C_{2n+1})$ , and equality holds since  $\alpha_{qc}(C_{2n+1}) \geq \alpha(C_{2n+1}) = n$ .

### 4.3 Links between the bounds $\gamma_r^{\text{col}}(G)$ , $\xi_r^{\text{col}}(G)$ , $\gamma_r^{\text{stab}}(G)$ , and $\xi_r^{\text{stab}}(G)$

In this last section, we make the link between the two hierarchies  $\{\xi_r^{\text{stab}}(G)\}$  (resp.  $\{\xi_r^{\text{col}}(G)\}$ ) and  $\{\gamma_r^{\text{stab}}(G)\}$  (resp.  $\{\gamma_r^{\text{col}}(G)\}$ ). The key tool is the interpretation of the coloring and stability numbers in terms of certain graph products.

We start with the (quantum) coloring number. For an integer  $k$ , recall that the Cartesian product  $G \square K_k$  of  $G$  and the complete graph  $K_k$  is the graph with vertex set  $V \times [k]$ , where two vertices  $(i, c)$  and  $(j, c')$  are adjacent if  $(\{i, j\} \in E$  and  $c = c')$  or  $(i = j$  and  $c \neq c')$ . The following is a well-known reduction of the chromatic number  $\chi(G)$  to the stability number of the Cartesian product  $G \square K_k$ :

$$\chi(G) = \min \{k \in \mathbb{N} : \alpha(G \square K_k) = |V|\}.$$

It was used in [19] to define the following lower bounds on the chromatic number:

$$\Lambda_r(G) = \min \{k \in \mathbb{N} : \text{las}_r^{\text{stab}}(G \square K_k) = |V|\},$$

where it was also shown that  $\text{las}_r^{\text{col}}(G) \leq \Lambda_r(G) \leq \chi(G)$  for all  $r \geq 1$ , with equality  $\Lambda_{|V|}(G) = \chi(G)$ . Hence the bounds  $\Lambda_r(G)$  may go beyond the fractional chromatic number. This is the case for the above mentioned Kneser graphs; see [18] for other graph instances.

The above reduction from coloring to stability number has been extended to the quantum setting in [32], where it is shown that

$$\chi_q(G) = \min \{k \in \mathbb{N} : \alpha_q(G \square K_k) = |V|\}.$$

It is therefore natural to use the upper bounds  $\xi_r^{\text{stab}}(G \square K_k)$  on  $\alpha_q(G \square K_k)$  in order to get the following lower bounds on the quantum coloring number:

$$\min \{k : \xi_r^{\text{stab}}(G \square K_k) = |V|\}, \quad (35)$$

which are thus the noncommutative analogues of the bounds  $\Lambda_r(G)$ .

Observe that, for any  $k \in \mathbb{N}$  and  $r \in \mathbb{N} \cup \{\infty, *\}$ , we have  $\xi_r^{\text{stab}}(G \square K_k) \leq |V|$ , which follows from Lemma 3 and the fact that the cliques  $C_i = \{(i, c) : c \in [k]\}$ , for  $i \in V$ , cover all vertices in  $G \square K_k$ . Let

$$\mathcal{C}_{G \square K_k} = \{g_{C_i} : i \in V\}, \quad \text{where } g_{C_i} = 1 - \sum_{c \in [k]} x_i^c,$$

denote the set of polynomials corresponding to these cliques. We now show that the parameter (35) in fact coincides with the parameter  $\gamma_r^{\text{col}}(G)$  for all  $r \in \mathbb{N} \cup \{\infty\}$ .

For this observe first that the quadratic polynomials in the set  $\mathcal{H}_{G,k}^{\text{col}}$  correspond precisely to the edges of  $G \square K_k$ , and that the projector constraints are included in  $\mathcal{I}_2(\mathcal{H}_{G,k}^{\text{col}})$  [see (28)]. Hence we have

$$\mathcal{I}_{2r}(\mathcal{H}_{G,k}^{\text{col}}) = \mathcal{I}_{2r}(\mathcal{H}_{G \square K_k} \cup \mathcal{C}_{G \square K_k}). \quad (36)$$

We will also use the following result.

**Lemma 4** *Let  $r \in \mathbb{N} \cup \{\infty, *\}$  and assume  $L$  is feasible for  $\xi_r^{\text{stab}}(G \square K_k)$ . Then, we have  $L(\sum_{i \in V, c \in [k]} x_i^c) = |V|$  if and only if  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{C}_{G \square K_k})$ .*

*Proof* Assume  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{C}_{G \square K_k})$ . Then  $0 = \sum_{i \in V} L(g_{C_i}) = |V| - L(\sum_{i,c} x_i^c)$ .

Conversely assume that  $0 = L(\sum_{i \in V, c \in [k]} x_i^c) - |V| = \sum_{i \in V} L(g_{C_i})$ . We will show  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{C}_{G \square K_k})$ . For this we first observe that  $g_{C_i} - (g_{C_i})^2 \in \mathcal{I}_2(\mathcal{H}_{G \square K_k})$  by (34). Hence  $L(g_{C_i}) = L(g_{C_i}^2) \geq 0$ , which, combined with  $\sum_i L(g_{C_i}) = 0$ , implies  $L(g_{C_i}) = 0$  for all  $i \in V$ . Next we show  $L(wg_{C_i}) = 0$  for all words  $w$  with degree at most  $2r - 1$ , using induction on  $\deg(w)$ . The base case  $w = 1$  holds by the above. Assume now  $w = uv$ , where  $\deg(v) < \deg(u) \leq r$ . Using the positivity of  $L$ , the

Cauchy-Schwarz inequality gives  $|L(uv g_{C_i})| \leq L(u^* u)^{1/2} L(v^* g_{C_i}^2 v)^{1/2}$ . Note that it suffices to show  $L(v^* g_{C_i} v) = 0$  since, using again (34), this implies  $L(v^* g_{C_i}^2 v) = 0$  and thus  $L(uv g_{C_i}) = 0$ . Using the tracial property of  $L$  and the induction assumption, we see that  $L(v^* g_{C_i} v) = L(vv^* g_{C_i}) = 0$  since  $\deg(vv^*) < \deg(w)$ .  $\square$

**Proposition 12** *For  $r \in \mathbb{N} \cup \{\infty\}$  we have  $\gamma_r^{\text{col}}(G) = \min\{k : \xi_r^{\text{stab}}(G \square K_k) = |V|\}$ .*

*Proof* Let  $L$  be a linear functional certifying  $\gamma_r^{\text{col}}(G) \leq k$ . Then, using (36) we see that  $L$  is feasible for  $\xi_r^{\text{stab}}(G \square K_k)$  and Lemma 4 shows that  $L(\sum_{i,c} x_i^c) = |V|$ . This shows  $\xi_r^{\text{stab}}(G \square K_k) \geq |V|$  and thus equality holds (since the reverse inequality always holds). Therefore,  $\min\{k : \xi_r^{\text{stab}}(G \square K_k) = |V|\} \leq k$ .

Conversely, assume  $\xi_r^{\text{stab}}(G \square K_k) = |V|$ . Since the optimum is attained, there exists a linear functional  $L$  feasible for  $\xi_r^{\text{stab}}(G \square K_k)$  with  $L(\sum_{i,c} x_i^c) = |V|$ . Using Lemma 4 we can conclude that  $L$  is zero on  $\mathcal{J}_{2r}(\mathcal{C}_{G \square K_k})$ . Hence, in view of (36),  $L$  is zero on  $\mathcal{J}_{2r}(\mathcal{H}_{G,k}^{\text{col}})$ . This shows  $\gamma_r^{\text{col}}(G) \leq k$ .  $\square$

Note that the proof of Proposition 12 also works in the commutative setting; this shows that the sequence  $\Lambda_r(G)$  corresponds to the usual Lasserre hierarchy for the feasibility problem defined by the equations (24)–(25), which is another way of showing  $\Lambda_\infty(G) = \chi(G)$ .

We now turn to the (quantum) stability number. For  $k \in \mathbb{N}$ , consider the graph product  $K_k \star G$ , with vertex set  $[k] \times G$ , and with an edge between two vertices  $(c, i)$  and  $(c', j)$  when  $(c \neq c', i = j)$  or  $(c = c', i \neq j)$  or  $(c \neq c', \{i, j\} \in E)$ . The product  $K_k \star G$  coincides with the homomorphic product  $K_k \ltimes \bar{G}$  used in [32, Sec. 4.2], where it is shown that

$$\alpha_q(G) = \max \{k \in \mathbb{N} : \alpha_q(K_k \star G) = k\}.$$

This suggests using the upper bounds  $\xi_r^{\text{stab}}(K_k \star G)$  on  $\alpha_q(K_k \star G)$  to define the following upper bounds on  $\alpha_q(G)$ :

$$\max \{k \in \mathbb{N} : \xi_r^{\text{stab}}(K_k \star G) = k\}. \quad (37)$$

For each  $c \in [k]$ , the set  $C^c = \{(c, i) : i \in V\}$  is a clique in  $K_k \star G$ , and we let

$$\mathcal{C}_{K_k \star G} = \{g_{C^c} : c \in [k]\}, \quad \text{where } g_{C^c} = 1 - \sum_{i \in V} x_c^i,$$

denote the set of polynomials corresponding to these cliques. As these  $k$  cliques cover the vertex set of  $K_k \star G$ , we can use Lemma 3 to conclude that  $\xi_r^{\text{stab}}(K_k \star G) \leq k$  for all  $r \in \mathbb{N} \cup \{\infty, *\}$ .

Again, observe that the quadratic polynomials in the set  $\mathcal{H}_{G,k}^{\text{stab}}$  correspond precisely to the edges of  $K_k \star G$  and that we have

$$\mathcal{J}_{2r}(\mathcal{H}_{G,k}^{\text{stab}}) = \mathcal{J}_{2r}(\mathcal{H}_{K_k \star G} \cup \mathcal{C}_{K_k \star G}).$$

Based on this, one can show the analogue of Lemma 4: If  $L$  is feasible for the program  $\xi_r^{\text{stab}}(K_k \star G)$ , then we have  $L(\sum_{i,c} x_c^i) = k$  if and only if  $L = 0$  on  $\mathcal{I}_{2r}(\mathcal{C}_{K_k \star G})$ . This lemma can be used to show the following result, whose proof is analogous to that of Proposition 12 and thus omitted.

**Proposition 13** *For  $r \in \mathbb{N} \cup \{\infty\}$  we have  $\gamma_r^{\text{stab}}(G) = \max\{k : \xi_r^{\text{stab}}(K_k \star G) = k\}$ .*

We do not know whether the results of Propositions 12 and 13 hold for  $r = *$ , because we do not know whether the supremum is attained in the program defining the parameter  $\xi_*^{\text{stab}}(\cdot) = \alpha_p(\cdot)$  (as was already observed in [51, p. 120]). Hence we can only claim the inequalities

$$\begin{aligned}\gamma_*^{\text{col}}(G) &\geq \min\{k : \xi_*^{\text{stab}}(G \square K_k) = |V|\} \quad \text{and} \\ \gamma_*^{\text{stab}}(G) &\leq \max\{k : \xi_*^{\text{stab}}(K_k \star G) = k\}.\end{aligned}$$

As mentioned above, we have  $\text{las}_r^{\text{col}}(G) \leq \Lambda_r(G)$  for any  $r \in \mathbb{N}$  [19, Prop. 3.3]. This result extends to the noncommutative setting and the analogous result holds for the stability parameters. In other words the hierarchies  $\{\gamma_r^{\text{col}}(G)\}$  and  $\{\gamma_r^{\text{stab}}(G)\}$  refine the hierarchies  $\{\xi_r^{\text{col}}(G)\}$  and  $\{\xi_r^{\text{stab}}(G)\}$ .

**Proposition 14** *For  $r \in \mathbb{N} \cup \{\infty, *\}$ ,  $\xi_r^{\text{col}}(G) \leq \gamma_r^{\text{col}}(G)$  and  $\xi_r^{\text{stab}}(G) \geq \gamma_r^{\text{stab}}(G)$ .*

*Proof* We may restrict to  $r \in \mathbb{N}$  since we have seen earlier that the inequalities hold for  $r \in \{\infty, *\}$ . The proof for the coloring parameters is similar to the proof of [19, Prop. 3.3] in the classical case and thus we omit it. We now show  $\xi_r^{\text{stab}}(G) \geq \gamma_r^{\text{stab}}(G)$ . Set  $k = \gamma_r^{\text{stab}}(G)$  and, using Proposition 13, let  $L \in \mathbb{R}\langle x_c^i : i \in V, c \in [k] \rangle_{2r}^*$  be optimal for  $\xi_r^{\text{stab}}(K_k \star G) = k$ . That is,  $L$  is tracial, symmetric, positive, and satisfies  $L(1) = 1$ ,  $L(\sum_{i,c} x_c^i) = k$ , and  $L = 0$  on  $\mathcal{I}(\mathcal{H}_{K_k \star G})$ . It suffices now to construct a tracial symmetric positive linear form  $\hat{L} \in \mathbb{R}\langle x_i : i \in V \rangle_{2r}^*$  such that  $\hat{L}(1) = 1$ ,  $\hat{L}(\sum_{i \in V} x_i) = k$ , and  $\hat{L} = 0$  on  $\mathcal{I}_{2r}(\mathcal{H}_G)$ , since this will imply  $\xi_r^{\text{stab}}(G) \geq k$ . For this, for any word  $x_{i_1}, \dots, x_{i_t}$  with degree  $1 \leq t \leq 2r$ , we define  $\hat{L}(x_{i_1}, \dots, x_{i_t}) := \sum_{c \in [k]} L(x_c^{i_1}, \dots, x_c^{i_t})$ , and we set  $\hat{L}(1) = L(1) = 1$ . Then, we have  $\hat{L}(\sum_{i \in V} x_i) = k$ . Moreover, one can easily check that  $\hat{L}$  is indeed tracial, symmetric, positive, and vanishes on  $\mathcal{I}_{2r}(\mathcal{H}_G)$ .  $\square$

**Acknowledgements** We thank the referees for their careful reading and useful suggestions.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Avis, D., Hasegawa, J., Kikuchi, Y., Sasaki, Y.: A quantum protocol to win the graph coloring game on all Hadamard graphs. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **E89–A**(5), 1378–1381 (2006)



2. Barker, G.P., Eifler, L.Q., Kezlan, T.P.: A non-commutative spectral theorem. *Linear Algebra Appl.* **20**(2), 95–100 (1978)
3. Bell, J.S.: On the Einstein Podolsky Rosen paradox. *Physics* **1**(3), 195–200 (1964)
4. Blackadar, B.: *Operator Algebras: Theory of  $C^*$ -Algebras and Von Neumann Algebras*. Encyclopaedia of Mathematical Sciences. Springer, New York (2006)
5. Brunner, N., Pironio, S., Acin, A., Gisin, N., Méthot, A.A., Scarani, V.: Testing the dimension of Hilbert spaces. *Phys. Rev. Lett.* **100**, 210,503 (2008)
6. Burgdorf, S., Cafuta, K., Klep, I., Povh, J.: The tracial moment problem and trace-optimization of polynomials. *Math. Program.* **137**(1), 557–578 (2013)
7. Burgdorf, S., Klep, I.: The truncated tracial moment problem. *J. Oper. Theory* **68**(1), 141–163 (2012)
8. Burgdorf, S., Klep, I., Povh, J.: *Optimization of Polynomials in Non-Commutative Variables*. Springer Briefs in Mathematics. Springer, New York (2016)
9. Cameron, P.J., Montanaro, A., Newman, M.W., Severini, S., Winter, A.: On the quantum chromatic number of a graph. *Electron. J. Comb.* **14**(1), 81 (2007)
10. Clauser, J.F., Horne, M.A., Shimony, A., Holt, R.A.: Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884 (1969)
11. Curto, R.E., Fialkow, L.A.: Solution of the Truncated Complex Moment Problem for Flat Data. In: *Memoirs of the American Mathematical Society*, vol. 568. American Mathematical Society (1996)
12. Doherty, A., Liang, Y.C., Toner, B., Wehner, S.: The quantum moment problem and bounds on entangled multiprover games. In: *Proceedings of the 2008 IEEE 23rd Annual Conference on Computational Complexity* pp. 199–210 (2008)
13. Dykema, K.J., Paulsen, V.I.: Synchronous correlation matrices and Connes’ embedding conjecture. *J. Math. Phys.* **57**, 015,214 (2016)
14. Dykema, K.J., Paulsen, V.I., Prakash, J.: Non-closure of the set of quantum correlations via graphs. [arXiv:1709.05032](https://arxiv.org/abs/1709.05032) (2017)
15. Fritz, T.: Tsirelson’s problem and Kirchberg’s conjecture. *Rev. Math. Phys.* **24**(5), 1250012 (2012)
16. Gribling, S., de Laat, D., Laurent, M.: Lower bounds on matrix factorization ranks via noncommutative polynomial optimization. [arXiv:1708.01573v2](https://arxiv.org/abs/1708.01573v2) (2017)
17. Gribling, S., de Laat, D., Laurent, M.: Matrices with high completely positive semidefinite rank. *Linear Algebra Appl.* **513**, 122–148 (2017)
18. Gvozdenović, N., Laurent, M.: Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization. *SIAM J. Optim.* **19**(2), 592–615 (2008)
19. Gvozdenović, N., Laurent, M.: The operator  $\psi$  for the chromatic number of a graph. *SIAM J. Optim.* **19**(2), 572–591 (2008)
20. Ji, Z.: Binary constraint system games and locally commutative reductions. [arXiv:1310.3794](https://arxiv.org/abs/1310.3794) (2013)
21. Junge, M., Navascués, M., Palazuelos, C., Perez-Garcia, D., Scholtz, V., Werner, R.: Connes’ embedding problem and Tsirelson’s problem. *J. Math. Phys.* **52**, 012,102 (2011)
22. Klep, I., Povh, J.: Constrained trace-optimization of polynomials in freely noncommuting variables. *J. Glob. Optim.* **64**(2), 325–348 (2016)
23. Klep, I., Schweighofer, M.: Connes’ embedding conjecture and sums of Hermitian squares. *Adv. Math.* **217**(4), 1816–1837 (2008)
24. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11**(3), 796–817 (2001)
25. Lasserre, J.B.: *Moments, Positive Polynomials and Their Applications*. Imperial College Press, London (2009)
26. Laurent, M.: A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0–1 programming. *Math. Oper. Res.* **28**(3), 470–496 (2003)
27. Laurent, M.: *Emerging Applications of Algebraic Geometry*. In: Putinar, M., Sullivant, S. (eds.) *Sums of Squares, Moment Matrices and Optimization Over Polynomials*, pp. 157–270. Springer, New York (2009)
28. Laurent, M., Piovesan, T.: Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone. *SIAM J. Optim.* **25**(4), 2461–2493 (2015)
29. Lovász, L.: Kneser’s conjecture, chromatic number, and homotopy. *J. Comb. Theory Ser A* **25**(3), 319–324 (1978)
30. Lovász, L.: On the shannon capacity of a graph. *IEEE Trans. Inf. Theory* **25**(1), 1–7 (2006)
31. Mančinska, L., Roberson, D.: Oddities of quantum colorings. *Baltic J. Mod. Comput.* **4**, 846–859 (2016)

32. Mančinska, L., Roberson, D.E.: Quantum homomorphisms. *J. Comb. Theory Ser. B* **118**, 228–267 (2016)
33. Mančinska, L., Scarpa, G., Severini, S.: New separations in zero-error channel capacity through projective Kochen–Specker sets and quantum coloring. *IEEE Trans. Inf. Theory* **59**(6), 4025–4032 (2013)
34. Mermin, N.D.: Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.* **65**, 3373–3376 (1990)
35. Navascués, M., Feix, A., Araujo, M., Vértesi, T.: Characterizing finite-dimensional quantum behavior. *Phys. Rev. A* **92**, 042117 (2015)
36. Navascués, M., Pironio, S., Acín, A.: A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.* **10**(7), 073,013 (2008)
37. Navascués, M., Pironio, S., Acín, A.: SDP relaxations for non-commutative polynomial optimization. In: Anjos M.F., Lasserre, J.B. (eds.) *Handbook on Semidefinite, Conic and Polynomial Optimization*, pp. 601–634. Springer (2012)
38. Navascués, M., Vértesi, T.: Bounding the set of finite dimensional quantum correlations. *Phys. Rev. Lett.* **115**(2), 020,501 (2015)
39. Nie, J.: Symmetric tensor nuclear norms. *SIAM J. Appl. Algebra Geom.* **1**(1), 599–625 (2017)
40. Nielsen, M.A., Chuang, I.L.: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge (2000)
41. Ortiz, C.M., Paulsen, V.I.: Quantum graph homomorphisms via operator systems. *Linear Algebra Appl.* **497**, 23–43 (2016)
42. Ozawa, N.: About the Connes’ embedding problem-algebraic approaches. *Jpn. J. Math.* **8**(1), 147–183 (2013)
43. Pál, K.F., Vértesi, T.: Efficiency of higher-dimensional Hilbert spaces for the violation of Bell inequalities. *Phys. Rev. A* **77**, 042,105 (2008)
44. Palazuelos, C., Vidick, T.: Survey on nonlocal games and operator space theory. *J. Math. Phys.* **57**(1), 015,220 (2016)
45. Parrilo, P.A.: *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, Caltech (2000)
46. Paulsen, V.I., Severini, S., Stahlke, D., Todorov, I.G., Winter, A.: Estimating quantum chromatic numbers. *J. Funct. Anal.* **270**(6), 2188–2222 (2016)
47. Peres, A.: Incompatible results of quantum measurements. *Phys. Lett. A* **151**(3), 107–108 (1990)
48. Pironio, S., Navascués, M., Acín, A.: Convergent relaxations of polynomial optimization problems with noncommuting variables. *SIAM J. Optim.* **20**(5), 2157–2180 (2010)
49. Prakash, A., Sikora, J., Varvitsiotis, A., Wei, Z.: Completely positive semidefinite rank. *Math. Program. Ser. A* (2017). <https://doi.org/10.1007/s10107-017-1198-4>
50. Prakash, A., Varvitsiotis, A.: Matrix factorizations of correlation matrices and applications. [arXiv:1702.06305](https://arxiv.org/abs/1702.06305) (2017)
51. Roberson, D.E.: *Variations on a Theme: Graph Homomorphisms*. PhD thesis, University of Waterloo (2013)
52. Sikora, J., Varvitsiotis, A.: Linear conic formulations for two-party correlations and values of nonlocal games. *Math. Program.* **162**(1), 431–463 (2017)
53. Sikora, J., Varvitsiotis, A., Wei, Z.: Minimum dimension of a Hilbert space needed to generate a quantum correlation. *Phys. Rev. Lett.* (2016)
54. Slofstra, W.: The set of quantum correlations is not closed. [arXiv:1703.08618](https://arxiv.org/abs/1703.08618) (2017)
55. Stark, C.: Learning optimal quantum models is NP-hard. [arXiv:1510.02800](https://arxiv.org/abs/1510.02800) (2015)
56. Szegedy, M.: A note on the theta number of Lovász and the generalized Delsarte bound. In: *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science*, pp. 36–39 (1994)
57. Tang, G., Shah, P.: Guaranteed tensor decomposition: a moment approach. In: *Proceedings of the 32nd International Conference on Machine Learning*, pp. 1491–1500 (2015)
58. Tsirelson, B.: Bell inequalities and operator algebras. Tech. rep. (2006). <http://www.tau.ac.il/~tsirel/download/bellopalg.pdf>
59. Wedderburn, J.H.M.: *Lectures on Matrices*. Dover Publications Inc., New York (1964)
60. Wehner, S., Christandl, M., Doherty, A.C.: Lower bound on the dimension of a quantum system given measured data. *Phys. Rev. A* **78**, 062,112 (2008)